GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: October 6, 2002, 12:40:30 ; Search time 284 Seconds
(without alignments)
120.909 Million cell updates/sec

Title: US-09-754-468-47
Perfect score: 20
Sequence: 1 gattagcataataaaatctc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1980.DAT:*
2: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1981.DAT:*
3: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1982.DAT:*
4: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1983.DAT:*
5: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1984.DAT:*
6: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1985.DAT:*
7: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1986.DAT:*
8: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1987.DAT:*
9: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1988.DAT:*
10: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1989.DAT:*
11: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1990.DAT:*
12: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1991.DAT:*
13: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1992.DAT:*
14: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1993.DAT:*
15: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1994.DAT:*
16: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1995.DAT:*
17: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1996.DAT:*
18: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1997.DAT:*
19: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1998.DAT:*
20: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA1999.DAT:*
21: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2000.DAT:*
22: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:*
23: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2001B.DAT:*
24: /net/abss06/SIDS1/gcgdata/hold-geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	20	100.0	836	22	AAH25706	E coli secA coding
	2	20	100.0	836	22	AAH25707	E coli secA mRNA a
c	3	20	100.0	2754	20	AAX88420	JP11192089 Seq ID
c	4	20	100.0	3811	20	AAX38291	E. coli secA DNA.
c	5	18.4	92.0	325	21	AAC05228	Human secreted pro
	6	18.4	92.0	645	22	AAK61816	Human immune/haema
	7	18.4	92.0	1972	22	AAK81074	Human immune/haema
	8	18.4	92.0	1972	22	AAK81075	Human immune/haema
	9	18.4	92.0	1976	22	AAK81076	Human immune/haema
c	10	17.4	87.0	3153	21	AAC59841	Human secreted pro
c	11	16.4	82.0	401	22	AAK95332	Human neuregulin g
c	12	16.4	82.0	401	22	AAK95333	Human neuregulin g
c	13	16.4	82.0	401	22	AAK96579	Human neuregulin g
c	14	16.4	82.0	401	22	AAK96581	Human neuregulin g
c	15	16.4	82.0	401	22	AAK96825	Human neuregulin g
c	16	16.4	82.0	401	22	AAK96826	Human neuregulin g
c	17	16.4	82.0	401	22	AAK98072	Human neuregulin g
c	18	16.4	82.0	401	22	AAK98074	Human neuregulin g
c	19	16.4	82.0	404	22	AAK96580	Human neuregulin g
c	20	16.4	82.0	404	22	AAK98073	Arabidopsis thalia
	21	16.4	82.0	589	21	AAC47899	Arabidopsis thalia
	22	16.4	82.0	610	21	AAC50288	Arabidopsis thalia
	23	16.4	82.0	663	21	AAC42090	Human secreted pro
	24	16.4	82.0	1589	22	AAS03908	cDNA coding human
c	25	16.4	82.0	3390	15	AAQ62283	cDNA coding human
c	26	16.4	82.0	3416	15	AAQ62282	Human foetal liver
c	27	15.8	79.0	570	22	ABA64122	Probe #12954 used
c	28	15.8	79.0	570	22	AAI44268	Human secreted pro
c	29	15.8	79.0	680	20	AAX20461	Human cDNA clone (
c	30	15.8	79.0	849	22	AAH04916	Human reverse tran
	31	15.8	79.0	1539	24	ABA92307	cDNA encoding a f
c	32	15.8	79.0	1605	20	AAX86281	Complement of cDNA
	33	15.8	79.0	1605	20	AAX86282	Flea epoxide hydro
c	34	15.8	79.0	1605	22	AAI66750	Flea epoxide hydro
	35	15.8	79.0	1605	22	AAI66751	Flea epoxide hydro
c	36	15.8	79.0	1605	22	AAC91967	Flea epoxide hydro
	37	15.8	79.0	1605	22	AAC91968	Flea epoxide hydro
c	38	15.8	79.0	1605	22	AAC86281	Flea epoxide hydro
	39	15.8	79.0	1605	22	AAC86282	Drosophila melanog
	40	15.8	79.0	4601	23	ABL09998	Human polynucleoti
c	41	15.8	79.0	5716	22	AAI61174	Human polynucleoti
c	42	15.8	79.0	5761	22	AAI59388	Tumour suppressor
c	43	15.8	79.0	8078	22	AAS46492	Human Goodpasture
c	44	15.8	79.0	12482	21	AAA50353	Human gene regulat
c	45	15.8	79.0	15853	24	AAS61457	

ALIGNMENTS

RESULT 1
AAH25706/c
ID AAH25706 standard; DNA; 836 BP.
XX
AC AAH25706;
XX
DT 14-AUG-2001 (first entry)
XX
DE E coli secA coding sequence.
XX
KW Antisense; microbial growth; essential gene; antimicrobial;
KW proliferation; infectious disease; secA; ds.
XX
OS Escherichia coli.
XX
PN US6228579-B1.
XX
PD 08-MAY-2001.
XX
PF 14-NOV-1997; 97US-0971090.
XX
PR 14-NOV-1997; 97US-0971090.
XX
PA (UYSA-) UNIV SAN DIEGO STATE FOUND.
XX
PI Zyskind JW, Forsyth RA;
XX
DR WPI; 2001-335011/35.
XX
PT Identifying microbial proliferation genes, useful for identifying
PT antimicrobial agents, comprises introducing into a microorganism an
PT exogenous nucleic acid having sequence identity to an endogenous

PT microbial gene -
XX
PS Disclosure; Fig 11; 28pp; English.
XX
CC The present invention describes a method of identifying genes essential
CC for microbial growth and proliferation, involving introducing an
CC exogenous nucleic acid into a microorganism, where the sequence is
CC similar to an endogenous microbial gene, and identifying the gene as
CC essential by comparing the organism's viability when the exogenous
CC sequence is expressed and when it is not present. This can be used to
CC identify targets for antimicrobial compounds for use in the therapy of
CC infectious diseases.
XX
SQ Sequence 836 BP; 223 A; 220 C; 198 G; 195 T; 0 other;

Query Match 100.0%; Score 20; DB 22; Length 836;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATTAGCATAATAAAATCTC 20
||||||||||||||||||||
Db 648 GATTAGCATAATAAAATCTC 629

RESULT 2
AAH25707
ID AAH25707 standard; RNA; 836 BP.
XX
AC AAH25707;
XX
DT 14-AUG-2001 (first entry)
XX
DE E coli secA mRNA antisense sequence.
XX
KW Antisense; microbial growth; essential gene; antimicrobial;
KW proliferation; infectious disease; secA; ss.
XX
OS Escherichia coli.
XX
PN US6228579-B1.
XX
PD 08-MAY-2001.
XX
PF 14-NOV-1997; 97US-0971090.
XX
PR 14-NOV-1997; 97US-0971090.
XX
PA (UYSA-) UNIV SAN DIEGO STATE FOUND.
XX
PI Zyskind JW, Forsyth RA;
XX
DR WPI; 2001-335011/35.
XX
PT Identifying microbial proliferation genes, useful for identifying
PT antimicrobial agents, comprises introducing into a microorganism an
PT exogenous nucleic acid having sequence identity to an endogenous
PT microbial gene -
XX
PS Example 3; Fig 12; 28pp; English.
XX
CC The present invention describes a method of identifying genes essential
CC for microbial growth and proliferation, involving introducing an
CC exogenous nucleic acid into a microorganism, where the sequence is
CC similar to an endogenous microbial gene, and identifying the gene as
CC essential by comparing the organism's viability when the exogenous
CC sequence is expressed and when it is not present. This can be used to
CC identify targets for antimicrobial compounds for use in the therapy of
CC infectious diseases.
XX
SQ Sequence 836 BP; 195 A; 198 C; 220 G; 223 U; 0 other;

Query Match 100.0%; Score 20; DB 22; Length 836;
Best Local Similarity 70.0%; Pred. No. 7.4;
Matches 14; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATTAGCATAATAAAATCTC 20
||::|||||:||:||||:|:|
Db 189 GAUUAGCAUAAUAAAAUCUC 208

RESULT 3
AAX88420/c
ID AAX88420 standard; DNA; 2754 BP.
XX
AC AAX88420;
XX
DT 30-SEP-1999 (first entry)
XX
DE JP11192089 Seq ID 4.
XX
KW Tp surface antigen; secretion-related enzyme; ds.
XX
OS Unidentified.
XX
PN JP11192089-A.
XX
PD 21-JUL-1999.
XX
PF 29-DEC-1997; 97JP-0367638.
XX
PR 29-DEC-1997; 97JP-0367638.
XX
PA (FJRE) FUJIREBIO KK.
XX
DR WPI; 1999-461459/39.
XX
PT Treponema pallidum-fused DNA sequence - and expression of Treponema
PT pallidum antigen by using said sequence
XX
PS Disclosure; Page 10-12; 16pp; Japanese.
XX
CC This invention describes a novel Treponema pallidum (Tp)-fused DNA
CC sequence in which a DNA sequence coding the surface antigen of Tp added
CC by a signal peptide is fused with a DNA sequence coding a
CC secretion-related enzyme. Also claimed is a method for expressing the
CC Tp antigen by using the above Tp-fused DNA sequence.
XX
SQ Sequence 2754 BP; 729 A; 680 C; 761 G; 584 T; 0 other;

Query Match 100.0%; Score 20; DB 20; Length 2754;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATTAGCATAATAAAATCTC 20
||||||||||||||||||||
Db 49 GATTAGCATAATAAAATCTC 30

RESULT 4
AAX38291/c
ID AAX38291 standard; DNA; 3811 BP.
XX
AC AAX38291;
XX
DT 16-JUN-1999 (first entry)
XX
DE E. coli secA DNA.
XX
KW Microorganism inhibitor; antisense; nuclease resistant; treatment;
KW ribonucleotide reductase; secA gene; pathological condition;
KW antimicrobial agent; crop protection; ss.
XX
OS Escherichia coli.
XX
PN WO9902673-A2.
XX

PD 21-JAN-1999.
XX
PF 10-JUL-1998; 98WO-CA00666.
XX
PR 10-JUL-1997; 97US-0052160.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Dugourd D, Wright JA, Young AH;
XX
DR WPI; 1999-120874/10.
XX
PT New oligonucleotides complementary to RR or SecA genes - useful to
PT inhibit growth of microorganisms
XX
PS Disclosure; Fig 5; 103pp; English.
XX
CC This invention describes novel antisense oligonucleotides
CC (AAX38301-X38552) which are nuclease resistant, and comprises about 3-50
CC nucleotides complementary to the ribonucleotide reductase gene or the
CC secA gene of a microorganism. The antisense oligonucleotides are used to
CC treat mammalian pathological conditions mediated by microorganisms. The
CC oligonucleotides are particularly useful as antimicrobial agents in crop
CC protection. This DNA sequence contains the Escherichia coli secA gene.
XX
SQ Sequence 3811 BP; 1016 A; 942 C; 1000 G; 853 T; 0 other;

Query Match 100.0%; Score 20; DB 20; Length 3811;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATTAGCATAATAAAATCTC 20
||||||||||||||||||||
Db 830 GATTAGCATAATAAAATCTC 811

RESULT 5
AAC05228/c
ID AAC05228 standard; cDNA; 325 BP.
XX
AC AAC05228;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 9303.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1; SEQ ID 9303; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 325 BP; 100 A; 68 C; 68 G; 89 T; 0 other;

Query Match 92.0%; Score 18.4; DB 21; Length 325;
Best Local Similarity 95.0%; Pred. No. 37;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GATTAGCATAATAAAATCTC 20
|| ||||||||||||||||||
Db 39 GAATAGCATAATAAAATCTC 20

RESULT 6
AAK61816
ID AAK61816 standard; cDNA; 645 BP.
XX
AC AAK61816;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:6876.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ss.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
DR P-PSDB; AAM89035.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS Claim 1; SEQ ID NO 6876; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 645 BP; 193 A; 145 C; 118 G; 186 T; 3 other;

Query Match 92.0%; Score 18.4; DB 22; Length 645;
Best Local Similarity 95.0%; Pred. No. 37;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 GATTAGCATAATAAAATCTC 20
          || ||||||||||||||||
Db     89 GAGTAGCATAATAAAATCTC 108


RESULT 7
AAK81074
ID   AAK81074 standard; DNA; 1972 BP.
XX
AC   AAK81074;
XX
DT   07-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:35886.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01354.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
```

PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS Disclosure; SEQ ID NO 35886; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 1972 BP; 660 A; 409 C; 350 G; 553 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 1972;
Best Local Similarity 95.0%; Pred. No. 37;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GATTAGCATAATAAAATCTC 20
|| ||||||||||||||||||
Db 89 GAGTAGCATAATAAAATCTC 108

RESULT 8
AAK81075
ID AAK81075 standard; DNA; 1972 BP.
XX
AC AAK81075;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:35887.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX

PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.

PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS Disclosure; SEQ ID NO 35887; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 1972 BP; 660 A; 409 C; 350 G; 553 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 1972;
Best Local Similarity 95.0%; Pred. No. 37;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GATTAGCATAATAAAATCTC 20
|| ||||||||||||||||
Db 89 GAGTAGCATAATAAAATCTC 108

RESULT 9
AAK81076
ID AAK81076 standard; DNA; 1976 BP.
XX
AC AAK81076;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:35888.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS Disclosure; SEQ ID NO 35888; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the

```
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention.
XX
SQ   Sequence 1976 BP; 663 A; 409 C; 350 G; 554 T; 0 other;

  Query Match            92.0%;  Score 18.4;  DB 22;  Length 1976;
  Best Local Similarity  95.0%;  Pred. No. 37;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GATTAGCATAATAAAATCTC 20
           || ||||||||||||||||||
Db      89 GAGTAGCATAATAAAATCTC 108


RESULT 10
AAC59841/c
ID   AAC59841 standard; DNA; 3153 BP.
XX
AC   AAC59841;
XX
DT   26-JAN-2001  (first entry)
XX
DE   Human secreted protein encoding DNA clone vo31 1.
XX
KW   Secreted protein; human; autoimmune disorder; multiple sclerosis; ulcer;
KW   systemic lupus erythematosus; rheumatoid arthritis; anaemia; stroke;
KW   haematopoiesis regulation; tissue regrowth; wound healing; haemophilia;
KW   Alzheimer's disease; Parkinson's disease; Shy-drager syndrome; cancer;
KW   contraceptive; infection; growth inhibition; hyperproliferative disorder;
KW   psoriasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200055375-A1.
XX
PD   21-SEP-2000.
XX
PF   17-MAR-2000; 2000WO-US07285.
XX
PR   17-MAR-1999;   99US-0124808.
PR   17-MAR-1999;   99US-0124916.
PR   17-AUG-1999;   99US-0149639.
PR   01-OCT-1999;   99US-0157247.
PR   29-NOV-1999;   99US-0167824.
PR   15-FEB-2000; 2000US-0182711.
XX
PA   (ALPH-) ALPHAGENE INC.
XX
PI   Valenzuela D,  Yuan O,  Hoffman H,  Hall J,  Rapiejko P;
XX
DR   WPI; 2000-638211/61.
DR   P-PSDB; AAB34740.
XX
PT   Novel proteins and polypeptides useful for the treatment of e.g
PT   multiple sclerosis, systemic lupus erythmatosus, rheumatoid arthritis,
PT   cancer, Alzheimer's disease, Parkinson's disease, stroke, anemia and
PT   ulcers -
XX
PS   Claim 116; Page 455-456; 493pp; English.
XX
CC   This invention relates to 59 human secreted proteins and the nucleotide
CC   sequences encoding them. Sequences AAC59788-C59846 and AAB34687-B34745
CC   represent the proteins and their encoding nucleotide sequences, and
CC   sequences AAB34746-B34771 represent fragments of the proteins. Probes
CC   for the DNA sequences are represented by sequences AAC59847-C59596. The
CC   proteins exhibit neuroprotective, dermatological, immunosuppressive,
CC   antiinflammatory, antianaemic, nootropic, antiparkinsonian,
CC   cerebroprotective, haemostatic, vulnerary, cytostatic, antipsoriatic,
CC   antibacterial, virucide, and fungicide activity. The proteins and
CC   nucleotide sequences are useful as nutritional sources or supplements
CC   and in research. The proteins are useful for treating immune deficiency
CC   and disorders, which may be genetic or resulting from infections,
CC   autoimmune disorders such as multiple sclerosis, systemic lupus
CC   erythmatosus, rheumatoid arthritis, and for treating myeloid or lymphoid
CC   cell deficiencies such as anaemias by regulating haematopoiesis. The
CC   proteins are also useful in compositions for bone, cartilage, tendon,
CC   ligament and/or nerve tissue growth or regeneration, for wound healing,
CC   tissue repair and replacement and in the treatment of wounds, incisions
CC   and ulcers. Other uses include in the treatment of central and
CC   peripheral nervous system and neuropathies such as Alzheimer's and
CC   Parkinson's diseases and Shy-Drager syndrome, and mechanical and
CC   traumatic disorders, such as spinal cord disorders, head trauma and
CC   stroke. The proteins may also be used as a contraceptive, and for
CC   treating coagulation disorders such as haemophilias. The protein and
CC   nucleotide sequences with cadherin activity are useful for treating
CC   cancer. Other uses for the protein include for inhibiting the growth,
CC   infection or function of, or killing, infectious agents such as bacteria,
CC   virus, fungi and other parasites, for effecting bodily characteristics
CC   such as height, weight, hair colour, effecting biorhythms or cardiac
CC   cycles or rhythms, effecting metabolism, catabolism, anabolism,
CC   processing, utilization, storage or elimination of dietary fat, lipid,
CC   protein, carbohydrate, vitamins, minerals, cofactors, effecting
CC   behavioural characteristics, providing analgesic effects and for treating
CC   hyperproliferative disorders such as psoriasis.
XX
SQ   Sequence 3153 BP; 938 A; 709 C; 647 G; 859 T; 0 other;

  Query Match            87.0%;  Score 17.4;  DB 21;  Length 3153;
  Best Local Similarity  94.7%;  Pred. No. 1e+02;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 ATTAGCATAATAAAATCTC 20
           ||| |||||||||||||||
Db    1944 ATTGGCATAATAAAATCTC 1926


RESULT 11
AAK95332/c
ID   AAK95332 standard; DNA; 401 BP.
XX
AC   AAK95332;
XX
DT   17-DEC-2001  (first entry)
XX
DE   Human neuregulin gene single nucleotide polymorphism SNP8NRG6787.
XX
KW   Human; neuregulin-1 associated gene 1; NRG1AG1; Schizophrenia gene;
KW   gene therapy; single nucleotide polymorphism; SNP; ds.
XX
OS   Homo sapiens.
XX
PN   WO200164876-A2.
XX
PD   07-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-US06376.
XX
PR   28-FEB-2000; 2000US-0515715.
XX
PA   (DECO-) DECODE GENETICS EHF.
XX
PI   Stefansson H,  Steinthorsdottir V,  Gulcher JR;
XX
DR   WPI; 2001-550179/61.
XX
PT   Neuregulin-1 associated gene 1 nucleic acids and fragments, useful for
PT   preventing diagnosing and treating schizophrenia -
XX
PS   Disclosure; Page 510; 750pp; English.
XX
CC   This sequence represents a single nucleotide polymorphism (SNP) of the
CC   human neuregulin-1 associated gene 1 (NRG1AG1) of the invention. The
```

CC NRG1AG1 gene is also referred to as the human Schizophrenia gene. The
CC invention also relates to fragments or variants of the gene and the
CC NRG1AG1 polypeptides they encode. The NRG1AG1 nucleic acids and
CC polypeptides may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate NRG1AG1 expression. For example,
CC they may be used to treat disorders associated with decreased expression
CC by rectifying mutations or deletions in a patient's genome that affect
CC the activity of NRG1AG1 by expressing inactive proteins or to supplement
CC the patients own production of NRG1AG1. Additionally, the gene may be
CC used to produce NRG1AG1 polypeptides, by inserting the nucleic acids into
CC a host cell and culturing the cell to express the protein. The gene may
CC also be used as DNA probes and primers in diagnostic assays to detect and
CC quantitate the presence of similar nucleic acids in samples, and
CC therefore which patients may be in need of restorative therapy. The
CC NRG1AG1 polypeptides may also be used as antigens in the production of
CC antibodies against NRG1AG1 and in assays to identify modulators of
CC NRG1AG1 expression and activity. Anti-NRG1AG1 antibodies and antagonists
CC may also be used to down regulate expression and activity. Anti-NRG1AG1
CC antibodies may also be used as diagnostic agents for detecting the
CC presence of NRG1AG1 polypeptides in samples. NRG1AG1 Is associated with
CC schizophrenia which may be prevented, diagnosed and/or treated by the
CC above methods.
XX
SQ Sequence 401 BP; 119 A; 96 C; 84 G; 98 T; 4 other;

Query Match 82.0%; Score 16.4; DB 22; Length 401;
Best Local Similarity 94.4%; Pred. No. 2.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GATTAGCATAATAAAATC 18
|||| ||||||||||||
Db 358 GATTTGCATAATAAAATC 341

RESULT 12
AAK95333/c
ID AAK95333 standard; DNA; 401 BP.
XX
AC AAK95333;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human neuregulin gene single nucleotide polymorphism SNP8NRG6844.
XX
KW Human; neuregulin-1 associated gene 1; NRG1AG1; Schizophrenia gene;
KW gene therapy; single nucleotide polymorphism; SNP; ds.
XX
OS Homo sapiens.
XX
PN WO200164876-A2.
XX
PD 07-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US06376.
XX
PR 28-FEB-2000; 2000US-0515715.
XX
PA (DECO-) DECODE GENETICS EHF.
XX
PI Stefansson H, Steinthorsdottir V, Gulcher JR;
XX
DR WPI; 2001-550179/61.
XX
PT Neuregulin-1 associated gene 1 nucleic acids and fragments, useful for
PT preventing diagnosing and treating schizophrenia -
XX
PS Disclosure; Page 510; 750pp; English.
XX
CC This sequence represents a single nucleotide polymorphism (SNP) of the
CC human neuregulin-1 associated gene 1 (NRG1AG1) of the invention. The
CC NRG1AG1 gene is also referred to as the human Schizophrenia gene. The
CC invention also relates to fragments or variants of the gene and the
CC NRG1AG1 polypeptides they encode. The NRG1AG1 nucleic acids and
CC polypeptides may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate NRG1AG1 expression. For example,
CC they may be used to treat disorders associated with decreased expression
CC by rectifying mutations or deletions in a patient's genome that affect
CC the activity of NRG1AG1 by expressing inactive proteins or to supplement
CC the patients own production of NRG1AG1. Additionally, the gene may be
CC used to produce NRG1AG1 polypeptides, by inserting the nucleic acids into
CC a host cell and culturing the cell to express the protein. The gene may
CC also be used as DNA probes and primers in diagnostic assays to detect and
CC quantitate the presence of similar nucleic acids in samples, and
CC therefore which patients may be in need of restorative therapy. The
CC NRG1AG1 polypeptides may also be used as antigens in the production of
CC antibodies against NRG1AG1 and in assays to identify modulators of
CC NRG1AG1 expression and activity. Anti-NRG1AG1 antibodies and antagonists
CC may also be used to down regulate expression and activity. Anti-NRG1AG1
CC antibodies may also be used as diagnostic agents for detecting the
CC presence of NRG1AG1 polypeptides in samples. NRG1AG1 Is associated with
CC schizophrenia which may be prevented, diagnosed and/or treated by the
CC above methods.
XX
SQ Sequence 401 BP; 114 A; 89 C; 92 G; 103 T; 3 other;

Query Match 82.0%; Score 16.4; DB 22; Length 401;
Best Local Similarity 94.4%; Pred. No. 2.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GATTAGCATAATAAAATC 18
|||| ||||||||||||
Db 301 GATTTGCATAATAAAATC 284

RESULT 13
AAK96579/c
ID AAK96579 standard; DNA; 401 BP.
XX
AC AAK96579;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human neuregulin gene insertion/deletion DNP8NRG1.
XX
KW Human; neuregulin-1 associated gene 1; NRG1AG1; Schizophrenia gene;
KW gene therapy; insertion; deletion; ds.
XX
OS Homo sapiens.
XX
PN WO200164876-A2.
XX
PD 07-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US06376.
XX
PR 28-FEB-2000; 2000US-0515715.
XX
PA (DECO-) DECODE GENETICS EHF.
XX
PI Stefansson H, Steinthorsdottir V, Gulcher JR;
XX
DR WPI; 2001-550179/61.
XX
PT Neuregulin-1 associated gene 1 nucleic acids and fragments, useful for
PT preventing diagnosing and treating schizophrenia -
XX
PS Disclosure; Page 729; 750pp; English.
XX
CC This sequence represents an insertion/deletion variant of the human
CC neuregulin-1 associated gene 1 (NRG1AG1) of the invention. The NRG1AG1
CC gene is also referred to as the human Schizophrenia gene. The invention
CC also relates to fragments or variants of the gene and the NRG1AG1
CC polypeptides they encode. The NRG1AG1 nucleic acids and polypeptides may
CC be used in the prevention, diagnosis and treatment of diseases associated
CC with inappropriate NRG1AG1 expression. For example, they may be used to
CC treat disorders associated with decreased expression by rectifying

CC mutations or deletions in a patient's genome that affect the activity of
CC NRG1AG1 by expressing inactive proteins or to supplement the patients own
CC production of NRG1AG1. Additionally, the gene may be used to produce
CC NRG1AG1 polypeptides, by inserting the nucleic acids into a host cell
CC and culturing the cell to express the protein. The gene may also be used
CC as DNA probes and primers in diagnostic assays to detect and quantitate
CC the presence of similar nucleic acids in samples, and therefore which
CC patients may be in need of restorative therapy. The NRG1AG1 polypeptides
CC may also be used as antigens in the production of antibodies against
CC NRG1AG1 and in assays to identify modulators of NRG1AG1 expression and
CC activity. Anti-NRG1AG1 antibodies and antagonists may also be used to
CC down regulate expression and activity. Anti-NRG1AG1 antibodies may
CC also be used as diagnostic agents for detecting the presence of NRG1AG1
CC polypeptides in samples. NRG1AG1 Is associated with schizophrenia which
CC may be prevented, diagnosed and/or treated by the above methods.
XX
SQ Sequence 401 BP; 124 A; 78 C; 80 G; 117 T; 2 other;

Query Match 82.0%; Score 16.4; DB 22; Length 401;
Best Local Similarity 94.4%; Pred. No. 2.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GATTAGCATAATAAAATC 18
|||| ||||||||||||
Db 186 GATTTGCATAATAAAATC 169

RESULT 14
AAK96581/c
ID AAK96581 standard; DNA; 401 BP.
XX
AC AAK96581;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human neuregulin gene insertion/deletion DNP8NRG3.
XX
KW Human; neuregulin-1 associated gene 1; NRG1AG1; Schizophrenia gene;
KW gene therapy; insertion; deletion; ds.
XX
OS Homo sapiens.
XX
PN WO200164876-A2.
XX
PD 07-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US06376.
XX
PR 28-FEB-2000; 2000US-0515715.
XX
PA (DECO-) DECODE GENETICS EHF.
XX
PI Stefansson H, Steinthorsdottir V, Gulcher JR;
XX
DR WPI; 2001-550179/61.
XX
PT Neuregulin-1 associated gene 1 nucleic acids and fragments, useful for
PT preventing diagnosing and treating schizophrenia -
XX
PS Disclosure; Page 729; 750pp; English.
XX
CC This sequence represents an insertion/deletion variant of the human
CC neuregulin-1 associated gene 1 (NRG1AG1) of the invention. The NRG1AG1
CC gene is also referred to as the human Schizophrenia gene. The invention
CC also relates to fragments or variants of the gene and the NRG1AG1
CC polypeptides they encode. The NRG1AG1 nucleic acids and polypeptides may
CC be used in the prevention, diagnosis and treatment of diseases associated
CC with inappropriate NRG1AG1 expression. For example, they may be used to
CC treat disorders associated with decreased expression by rectifying
CC mutations or deletions in a patient's genome that affect the activity of
CC NRG1AG1 by expressing inactive proteins or to supplement the patients own
CC production of NRG1AG1. Additionally, the gene may be used to produce
CC NRG1AG1 polypeptides, by inserting the nucleic acids into a host cell
CC and culturing the cell to express the protein. The gene may also be used
CC as DNA probes and primers in diagnostic assays to detect and quantitate
CC the presence of similar nucleic acids in samples, and therefore which
CC patients may be in need of restorative therapy. The NRG1AG1 polypeptides
CC may also be used as antigens in the production of antibodies against
CC NRG1AG1 and in assays to identify modulators of NRG1AG1 expression and
CC activity. Anti-NRG1AG1 antibodies and antagonists may also be used to
CC down regulate expression and activity. Anti-NRG1AG1 antibodies may
CC also be used as diagnostic agents for detecting the presence of NRG1AG1
CC polypeptides in samples. NRG1AG1 Is associated with schizophrenia which
CC may be prevented, diagnosed and/or treated by the above methods.
XX
SQ Sequence 401 BP; 119 A; 74 C; 85 G; 123 T; 0 other;

Query Match 82.0%; Score 16.4; DB 22; Length 401;
Best Local Similarity 94.4%; Pred. No. 2.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GATTAGCATAATAAAATC 18
|||| ||||||||||||
Db 90 GATTTGCATAATAAAATC 73

RESULT 15
AAK96825/c
ID AAK96825 standard; DNA; 401 BP.
XX
AC AAK96825;
XX
DT 17-DEC-2001 (first entry)
XX
DE Human neuregulin gene single nucleotide polymorphism SNP8NRG6787.
XX
KW Human; neuregulin 1 gene; schizophrenia; gene therapy; SNP;
KW single nucleotide polymorphism; ds.
XX
OS Homo sapiens.
XX
PN WO200164877-A2.
XX
PD 07-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US06377.
XX
PR 28-FEB-2000; 2000US-0515716.
XX
PA (DECO-) DECODE GENETICS EHF.
XX
PI Stefansson H, Steinthorsdottir V, Gulcher JR;
XX
DR WPI; 2001-514841/56.
XX
PT Neuregulin 1 nucleic acids and proteins useful for diagnosing
PT preventing and treating schizophrenia -
XX
PS Disclosure; Page 95; 756pp; English.
XX
CC This sequence represents a single nucleotide polymorphism (SNP)
CC from the human neuregulin 1 gene of the invention.
CC The invention also relates to fragments or variants of the neuregulin 1
CC gene. The gene and its proteins may be used in the prevention, diagnosis
CC and treatment of diseases associated with inappropriate neuregulin 1
CC expression, such as schizophrenia. For example they may be used to treat
CC disorders associated with decreased neuregulin 1 expression by rectifying
CC mutations or deletions in a patient's genome that affect the activity of
CC neuregulin 1 by expressing inactive proteins or to supplement the
CC patients own production of polypeptides. Additionally, the gene may be
CC used to produce the neuregulin 1 protein, by inserting the nucleic acids
CC into a host cell and culturing the cell to express the protein. The gene
CC and its complementary sequences may also be used as DNA probes in
CC diagnostic assays to detect and quantitate the presence of similar
CC nucleic acids in samples, and therefore which patients may be in need of
CC restorative therapy. The protein may also be used as antigens in the

CC production of antibodies against neuregulin 1 and in assays to identify
CC modulators of neuregulin 1 expression and activity. The antibodies and
CC antagonists may also be used to down regulate expression and activity.
CC The antibodies may also be used as diagnostic agents for detecting the
CC presence of neuregulin 1 in samples.
XX
SQ Sequence 401 BP; 119 A; 96 C; 84 G; 98 T; 4 other;

Query Match 82.0%; Score 16.4; DB 22; Length 401;
Best Local Similarity 94.4%; Pred. No. 2.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 GATTAGCATAATAAAATC 18
           |||| ||||||||||||||
Db     358 GATTTGCATAATAAAATC 341
```

Search completed: October 6, 2002, 15:21:06
Job time : 290 secs

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: October 6, 2002, 13:13:14 ; Search time 2261 Seconds
(without alignments)
185.109 Million cell updates/sec

Title: US-09-754-468-47
Perfect score: 20
Sequence: 1 gattagcataataaaatctc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	100.0	20	6	AX191765	AX191765 Sequence
c 2	20	100.0	836	6	AR149152	AR149152 Sequence
3	20	100.0	836	6	AR149153	AR149153 Sequence
c 4	20	100.0	2048	1	ECOENVAA	M19211 E.coli cell
c 5	20	100.0	2754	6	E27365	E27365 Treponema p
c 6	20	100.0	3811	1	ECOSECA	M20791 Escherichia
c 7	20	100.0	12434	1	AE000119	AE000119 Escherich
c 8	20	100.0	12518	1	AE005186	AE005186 Escherich
c 9	20	100.0	28277	1	EC2MIN	X55034 E. coli 2 m
c 10	20	100.0	28277	6	AX191720	AX191720 Sequence
c 11	20	100.0	111401	1	ECO110K	D10483 E.coli K12
12	19	95.0	155976	2	AC055113	AC055113 Homo sapi
c 13	19	95.0	179212	9	AC084877	AC084877 Homo sapi
c 14	18.4	92.0	22286	1	AE008700	AE008700 Salmonell
c 15	18.4	92.0	127603	9	AC036102	AC036102 Homo sapi
c 16	18.4	92.0	138846	2	AC068478	AC068478 Homo sapi
17	18.4	92.0	156978	2	AC023379	AC023379 Homo sapi
c 18	18.4	92.0	157385	9	AC006023	AC006023 Homo sapi
19	18.4	92.0	165102	2	AC074279	AC074279 Homo sapi
c 20	18.4	92.0	166427	2	AC069593	AC069593 Homo sapi
21	18.4	92.0	168004	2	AC019203	AC019203 Homo sapi
22	18.4	92.0	172579	9	AC008064	AC008064 Homo sapi
c 23	18.4	92.0	183084	2	AC025936	AC025936 Homo sapi
c 24	18.4	92.0	188049	2	AC069391	AC069391 Homo sapi
25	18.4	92.0	193588	2	AC025370	AC025370 Homo sapi
26	18.4	92.0	194782	2	AC092433	AC092433 Homo sapi
27	18.4	92.0	200475	9	AC079271	AC079271 Homo sapi
c 28	18.4	92.0	201629	9	AL158047	AL158047 Human DNA
c 29	18.4	92.0	204790	2	AC013452	AC013452 Homo sapi
c 30	18.4	92.0	251050	1	AL627265	AL627265 Salmonell
c 31	18.4	92.0	347253	9	AF363578	AF363578 Homo sapi
32	18	90.0	51869	2	AC094868	AC094868 Rattus no
33	18	90.0	175456	2	AL359976	AL359976 Homo sapi
c 34	18	90.0	195290	2	AC074364	AC074364 Homo sapi
c 35	17.4	87.0	942	33	AC038729	Ac038729 Giardia i
c 36	17.4	87.0	954	33	AC067176	Ac067176 Giardia i
37	17.4	87.0	63769	2	AC079334	AC079334 Homo sapi
38	17.4	87.0	73610	9	AL359735	AL359735 Human DNA
c 39	17.4	87.0	102990	2	AP000801	AP000801 Homo sapi
40	17.4	87.0	109568	2	AP000840	AP000840 Homo sapi
41	17.4	87.0	126755	9	HSDJ18C9	AL049709 Human DNA
c 42	17.4	87.0	138969	9	AC060771	AC060771 Homo sapi
43	17.4	87.0	142278	9	AC005177	AC005177 Homo sapi
c 44	17.4	87.0	150406	2	AC098684	AC098684 Mus muscu
45	17.4	87.0	161742	2	AC036204	AC036204 Homo sapi

ALIGNMENTS

RESULT 1
AX191765
LOCUS AX191765 20 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 47 from Patent WO0149775.
ACCESSION AX191765
VERSION AX191765.1 GI:15209934
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 1 to 20)
AUTHORS Iversen,P.L.
TITLE Antisense antibacterial cell division composition and method
JOURNAL Patent: WO 0149775-A 47 12-JUL-2001;
Avi Biopharma, Inc. (US)
FEATURES Location/Qualifiers
source 1. .20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="oligonucleotide"
BASE COUNT 9 a 3 c 2 g 6 t
ORIGIN

TOO LATE SAME INV

```
  Query Match            100.0%;  Score 20;  DB 6;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 2.2e+02;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GATTAGCATAATAAAATCTC 20
          ||||||||||||||||||||
Db      1 GATTAGCATAATAAAATCTC 20


RESULT 2
AR149152/c
LOCUS       AR149152                836 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 6 from patent US 6228579.
ACCESSION   AR149152
VERSION     AR149152.1  GI:15113743
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 836)
  AUTHORS   Zyskind,J.W. and Forsyth,R.Allyn.
  TITLE     Method for identifying microbial proliferation genes
  JOURNAL   Patent: US 6228579-A 6 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .836
                     /organism="unknown"
BASE COUNT      223 a    220 c    198 g    195 t
ORIGIN

  Query Match            100.0%;  Score 20;  DB 6;  Length 836;
  Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GATTAGCATAATAAAATCTC 20
          ||||||||||||||||||||
Db    648 GATTAGCATAATAAAATCTC 629


RESULT 3
AR149153
LOCUS       AR149153                836 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 7 from patent US 6228579.
ACCESSION   AR149153
VERSION     AR149153.1  GI:15113744
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 836)
  AUTHORS   Zyskind,J.W. and Forsyth,R.Allyn.
  TITLE     Method for identifying microbial proliferation genes
  JOURNAL   Patent: US 6228579-A 7 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .836
                     /organism="unknown"
BASE COUNT      195 a    198 c    220 g    223 t
ORIGIN

  Query Match            100.0%;  Score 20;  DB 6;  Length 836;
  Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GATTAGCATAATAAAATCTC 20
          ||||||||||||||||||||
Db    189 GATTAGCATAATAAAATCTC 208


RESULT 4
ECOENVAA/c
LOCUS       ECOENVAA                2048 bp ss-DNA     linear   BCT 20-DEC-1995
DEFINITION  E.coli cell permeability-cell separation protein (envA) gene,
            complete cds., ftsZ gene, 3' end, and secA gene, 5' end.
ACCESSION   M19211
VERSION     M19211.1  GI:145846
KEYWORDS    cell permeability-cell separation protein; envA gene; ftsZ gene;
            secA gene.
SOURCE      Escherichia coli (strain K-12) (clone:  pACYC184.) DNA.
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
            Escherichia.
REFERENCE   1  (bases 1 to 2048)
  AUTHORS   Beall,B. and Lutkenhaus,J.
  TITLE     Sequence analysis, transcriptional organization, and insertional
            mutagenesis of the envA gene of Escherichia coli
  JOURNAL   J. Bacteriol. 169 (12), 5408-5415 (1987)
  MEDLINE   88058745
COMMENT     Draft entry and computer readable of sequence [1] kindly provided
            by J.Lutkenhaus (19-FEB-1988). (put. see Kay for more info.).
FEATURES             Location/Qualifiers
     source          1. .2048
                     /organism="Escherichia coli"
                     /strain="K-12"
                     /db_xref="taxon:562"
                     /clone="pACYC184."
     gene            1. .24
                     /gene="ftsZ"
     CDS             <1. .24
                     /gene="ftsZ"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="AAA83848.1"
                     /db_xref="GI:145847"
                     /translation="FLRKQAD"
     gene            125. .1042
                     /gene="envA"
     CDS             125. .1042
                     /gene="envA"
                     /codon_start=1
                     /transl_table=11
                     /product="ell permeability-cell separation protein"
                     /protein_id="AAA83849.1"
                     /db_xref="GI:145848"
                     /translation="MIKQRTLKRIVQATGVGLHTGKKVTLTLRPAPANTGVIYRRTDL
                     NPPVDFPADAKSVRDTMLCTCLVNEHDVRISTVEHLNAALAGLGIDNIVIEVNAPEIP
                     IMDGSAAPFVYLLLDAGIDELNCAKKFVRIKETVRVEDGDKWAEFKPYNGFSLDFTID
                     FNHPAIDSSNQRYAMNFSADAFMRQISRARTFGFMRDIEYLQSRGLCLGGSFDCAIVV
                     DDYRVLNEDGLRFEDEFVRHKMLDAIGDLFMCGHNIIGAFTAYKSGHALNNKLLQAVL
                     AKQEAWEYVTFQDDAELPLAFKAPSAVLA"
     CDS             1342. .1785
                     /note="ORF; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="unknown protein"
                     /protein_id="AAA83850.1"
                     /db_xref="GI:1128950"
                     /translation="MVAASLGLPALSNAAEPNAPAKATTRNHEPSAKVNFGQLALLEA
                     NTRRPNSNYSVDYWHQHAIRTVIRHLSFAMAPQTLPVAEESLPLQAQHLALLDTLSAL
                     LTQEGTPSEKGYRIDYAHFTPQAKFSTPVWISQAQGIRAGPQRLT"
     gene            1847. .2048
                     /gene="secA"
     CDS             1847. .>2048
                     /gene="secA"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="AAA83851.1"
                     /db_xref="GI:551802"
                     /translation="MLIKLLTKVFGSRNDRTLRRMRKVVNIINAMEPEMEKLSDEELK
                     GKTAEFRARLEKGEVLENLIPEA"
BASE COUNT      509 a    510 c    499 g    530 t
ORIGIN      390 bp upstream of ClaI site.

  Query Match            100.0%;  Score 20;  DB 1;  Length 2048;
  Best Local Similarity  100.0%;  Pred. No. 88;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GATTAGCATAATAAAATCTC 20
```

```
          ||||||||||||||||||||
Db    1855 GATTAGCATAATAAAATCTC 1836


RESULT 5
E27365/c
LOCUS       E27365                2754 bp    DNA     linear   PAT 07-FEB-2001
DEFINITION  Treponema pallidum-fused DNA sequence and method for expressing T.
            pallidum antigen with the use of the said sequence.
ACCESSION   E27365
VERSION     E27365.1  GI:13018177
KEYWORDS    JP 1999192089-A/4.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 2754)
  AUTHORS   Katsuya,F.Y.H.H. and Ito.
  TITLE     Treponema pallidum-fused DNA sequence and method for expressing T.
            pallidum antigen with the use of the said sequence
  JOURNAL   Patent: JP 1999192089-A 4 21-JUL-1999;
            FUJIREBIO INC
COMMENT     OS   Unidentified
            PN   JP 1999192089-A/4
            PD   21-JUL-1999
            PF   29-DEC-1997 JP 1997367638
            PR
            PI   KATSUYA FUJIMURA,YASUHIRO HARA,SATOSHI ITO
            PC   C12N15/09//C07K14/20,C07K19/00,C12N1/21,C12P21/02,G01N33/53,
            PC   G01N33/577,
            PC   (C12N15/09,C12R1:01),(C12N1/21,C12R1:19),(C12P21/02,C12R1:19),
            PC   C12N15/00,
            PC   (C12N15/00,C12R1:01)
            CC   Strandedness: Double;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .2754
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .2754
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT      729 a    680 c    761 g    584 t
ORIGIN

  Query Match             100.0%;  Score 20;  DB 6;  Length 2754;
  Best Local Similarity   100.0%;  Pred. No. 83;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GATTAGCATAATAAAATCTC 20
          ||||||||||||||||||||
Db     49 GATTAGCATAATAAAATCTC 30


RESULT 6
ECOSECA/c
LOCUS       ECOSECA               3811 bp    DNA     linear   BCT 26-APR-1993
DEFINITION  Escherichia coli SecA protein gene, complete cds; mutT 5' end.
ACCESSION   M20791
VERSION     M20791.1  GI:147792
KEYWORDS    secA protein.
SOURCE      E.coli (strain MC4100) DNA.
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
            Escherichia.
REFERENCE   1  (bases 1 to 3811)
  AUTHORS   Schmidt,M., Rollo,E., Grodberg,I. and Oliver,D.
  TITLE     Nucleotide sequence of secA gene and secA(ts) mutations preventing
            protein export in Escherichia coli
  JOURNAL   J. Bacteriol. 170, 3404-3414 (1988)
  MEDLINE   88298644
COMMENT     Draft entry and computer readable sequence [1] kindly submitted by
            M. Schmidt 28-SEPT-1988
            The mutT gene was identified in Mol. Gen. Genet. 206, 9-16 (1987)
            accession number Xo4831.
FEATURES             Location/Qualifiers
     source          1. .3811
                     /organism="Escherichia coli"
                     /db_xref="taxon:562"
     CDS             317. .760
                     /note="gene X peptide"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="AAA24618.1"
                     /db_xref="GI:147793"
                     /translation="MVAASLGLPALSNAAEPNAPAKATTRNHEPSAKVNFGQLALLEA
                     NTRRPNSNYSVDYWHQHAIRTVIRHLSFAMAPQTLPVAEESLPLQAQHLALLDTLSAL
                     LTQEGTPSEKGYRIDYAHFTPQAKFSTPVWISQAQGIRAGPQRLT"
     gene            822. .3527
                     /gene="secA"
     CDS             822. .3527
                     /gene="secA"
                     /note="secA protein"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="AAA24619.1"
                     /db_xref="GI:147794"
                     /translation="MLIKLLTKVFGSRNDRTLRRMRKVVNIINAMEPEMEKLSDEELK
                     GKTAEFRARLEKGEVLENLIPEAFAVVREASKRVFGMRHFDVQLLGGMVLNERCIAEM
                     RTGEGKTLTATLPAYLNALTGKGVHVVTVNDYLAQRDAENNRPLFEFLGLTVGINLPG
                     MPAPAKREAYAADITYGTNNEYGFDYLRDNMAFSPEERVQRKLHYALVDEVDSILIDE
                     ARTPLIISGPAEDSSEMYKRVNKIIPHLIRQEKEDSETFQGEGHFSVDEKSRQVNLTE
                     RGLVLIEELLVKEGIMDEGESLYSPANIMLMHHVTAALRAHALFTRDVDYIVKDGEVI
                     IVDEHTGRTMQGRRWSDGLHQAVEAKEGVQIQNENQTLASITFQNYFRLYEKLAGMTG
                     TADTEAFEFSSIYKLDTVVVPTNRPMIRKDLPDLVYMTEAEKIQAIIEDIKERTAKGQ
                     PVLVGTISIEKSELVSNELTKAGIKHNVLNAKFHANEAAIVAQAGYPAAVTIATNMAG
                     RGTDIVLGGSWQAEVAALENPTAEQIEKIKADWQVRHDAVLEAGGLHIIGTERHESRR
                     IDNQLRGRSGRQGDAGSSRFYLSMEDALMRIFASDRVSGMMRKLGMKPGEAIEHPWVT
                     KAIANAQRKVESRNFDIRKQLLEYDDVANDQRRAIYSQRNELLDVSDVSETINSIRED
                     VFKATIDAYIPPQSLEEMWDIPGLQERLKNDFDLDLPIAEWLDKEPELHEETLRDGIL
                     AQSIEVYQRKEEVVGAEMMRHFEKGVMLQTLDSLWKEHLAAMDYLRQGIHLRGYAQKD
                     PKQEYKRESFSMFAAMLESLKYEVISTLSKVQVRMPEEVEELEQQRRMEAERLAQMQQ
                     LSHQDDDSAAAAALAAQTGERKVGRNDPCPCGSGKKYKQCHGRLQ"
     gene            3587. .3811
                     /gene="mutT"
     CDS             3587. .>3811
                     /gene="mutT"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="AAA24620.1"
                     /db_xref="GI:147795"
                     /translation="MKKLQIAVGIIRNENNEIFITRRAADAHMANKLEFPGGKIEMGE
                     TPEQAVVRELQEEVGITPQHFSLFEKLEYEF"
BASE COUNT     1016 a    938 c   1004 g    853 t
ORIGIN      3 bp upstream of PvuII site.

  Query Match             100.0%;  Score 20;  DB 1;  Length 3811;
  Best Local Similarity   100.0%;  Pred. No. 78;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GATTAGCATAATAAAATCTC 20
          ||||||||||||||||||||
Db    830 GATTAGCATAATAAAATCTC 811


RESULT 7
AE000119/c
LOCUS       AE000119              12434 bp    DNA     linear   BCT 01-DEC-2000
DEFINITION  Escherichia coli K12 MG1655 section 9 of 400 of the complete
            genome.
ACCESSION   AE000119 U00096
VERSION     AE000119.1  GI:1786283
KEYWORDS    .
SOURCE      Escherichia coli K12.
  ORGANISM  Escherichia coli K12
            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
            Escherichia.
```

```
REFERENCE   1  (bases 1 to 12434)
  AUTHORS   Blattner,F.R., Plunkett,G. III, Bloch,C.A., Perna,N.T., Burland,V.,
            Riley,M., Collado-Vides,J., Glasner,J.D., Rode,C.K., Mayhew,G.F.,
            Gregor,J., Davis,N.W., Kirkpatrick,H.A., Goeden,M.A., Rose,D.J.,
            Mau,B. and Shao,Y.
  TITLE     The complete genome sequence of Escherichia coli K-12
  JOURNAL   Science 277 (5331), 1453-1474 (1997)
  MEDLINE   97426617
   PUBMED   9278503
REFERENCE   2  (bases 1 to 12434)
  AUTHORS   Blattner,F.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics,
            University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
            Email: ecoli@genetics.wisc.edu Phone: 608-262-2534 Fax:
            608-263-7459
REFERENCE   3  (bases 1 to 12434)
  AUTHORS   Blattner,F.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics,
            University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
            Email: ecoli@genetics.wisc.edu Phone: 608-262-2534 Fax:
            608-263-7459
REFERENCE   4  (bases 1 to 12434)
  AUTHORS   Plunkett,G. III.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-OCT-1998) Laboratory of Genetics, University of
            Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
COMMENT     This sequence was determined by the E. coli Genome Project at the
            University of Wisconsin-Madison (Frederick R. Blattner, director).
            Supported by NIH grants HG00301 and HG01428 (from the Human Genome
            Project and NCHGR). The entire sequence was independently
            determined from E. coli K12 strain MG1655. Predicted open reading
            frames were determined using GeneMark software, kindly supplied by
            Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA,
            30332 [e-mail: mark@amber.gatech.edu]. Open reading frames that
            have been correlated with genetic loci are being annotated with CG
            Site Nos., unique ID nos. for the genes in the E. coli Genetic
            Stock Center (CGSC) database at Yale University, kindly supplied by
            Mary Berlyn. A public version of the database is accessible
            (http://cgsc.biology.yale.edu). Annotation of the genome is an
            ongoing task whose goal is to make the genome sequence more useful
            by correlating it with other data.  Comments to the authors are
            appreciated. Updated information will be available at the E. coli
            Genome Project's World Wide Web site
            (http://www.genetics.wisc.edu). *** The E. coli K12 sequence and
            its annotations are periodically updated; this is version M54. No
            sequence changes. Annotation updates: updated gene identifications
            and products; all new functional assignments courtesy of Monica
            Riley; added promoters, protein binding sites, and repeated
            sequences described in reference 1. The unique numeric identifiers
            beginning with a lowercase 'b' assigned to each gene (protein- or
            RNA-encoding) are now designated as gene synonyms instead of
            labels. This should allow them to be searched for in Entrez as gene
            names.
FEATURES             Location/Qualifiers
     source          1. .12434
                     /organism="Escherichia coli K12"
                     /strain="K12"
                     /sub_strain="MG1655"
                     /db_xref="taxon:83333"
     gene            61. .1212
                     /gene="ftsZ"
                     /note="b0095"
     CDS             61. .1212
                     /gene="ftsZ"
                     /function="enzyme; Cell division"
                     /note="o383; 99 pct identical to FTSZ_ECOLI SW: P06138"
                     /codon_start=1
                     /transl_table=11
                     /product="cell division; forms circumferential ring;
                     tubulin-like GTP-binding protein and GTPase"
                     /protein_id="AAC73206.1"
                     /db_xref="GI:1786284"
                     /translation="MFEPMELTNDAVIKVIGVGGGGGNAVEHMVRERIEGVEFFAVNT
                     DAQALRKTAVGQTIQIGSGITKGLGAGANPEVGRNAADEDRDALRAALEGADMVFIAA
                     GMGGGTGTGAAPVVAEVAKDLGILTVAVVTKPFNFEGKKRMAFAEQGITELSKHVDSL
                     ITIPNDKLLKVLGRGISLLDAFGAANDVLKGAVQGIAELITRPGLMNVDFADVRTVMS
                     EMGYAMMGSGVASGEDRAEEAAEMAISSPLLEDIDLSGARGVLVNITAGFDLRLDEFE
                     TVGNTIRAFASDNATVVIGTSLDPDMNDELRVTVVATGIGMDKRPEITLVTNKQVQQP
                     VMDRYQQHGMAPLTQEQKPVAKVVNDNAPQTAKEPDYLDIPAFLRKQAD"
     promoter        1245. .1272
                     /note="factor Sigma70; predicted +1 start at 106523"
     gene            1313. .2230
                     /gene="lpxC"
                     /note="b0096"
     CDS             1313. .2230
                     /gene="lpxC"
                     /function="enzyme; Surface polysaccharides and antigens"
                     /note="o305; 100 pct identical to LPXC_ECOLI SW: P07652;
                     alternate gene name envA"
                     /codon_start=1
                     /transl_table=11
                     /product="UDP-3-O-acyl N-acetylglucosamine deacetylase;
                     lipid A biosynthesis"
                     /protein_id="AAC73207.1"
                     /db_xref="GI:1786285"
                     /translation="MIKQRTLKRIVQATGVGLHTGKKVTLTLRPAPANTGVIYRRTDL
                     NPPVDFPADAKSVRDTMLCTCLVNEHDVRISTVEHLNAALAGLGIDNIVIEVNAPEIP
                     IMDGSAAPFVYLLLDAGIDELNCAKKFVRIKETVRVEDGDKWAEFKPYNGFSLDFTID
                     FNHPAIDSSNQRYAMNFSADAFMRQISRARTFGFMRDIEYLQSRGLCLGGSFDCAIVV
                     DDYRVLNEDGLRFEDEFVRHKMLDAIGDLFMCGHNIIGAFTAYKSGHALNNKLLQAVL
                     AKQEAWEYVTFQDDAELPLAFKAPSAVLA"
     promoter        2300. .2328
                     /note="factor Sigma70; predicted +1 start at 107579"
     promoter        2348. .2378
                     /note="factor Sigma70; predicted +1 start at 107629"
     gene            2386. .2973
                     /gene="yacA"
                     /note="b0097"
     CDS             2386. .2973
                     /gene="yacA"
                     /function="orf; Unknown"
                     /note="o195; 100 pct identical to YACA_ECOLI SW:
                     P10409(147aa) but has 48 additional N-term residues; GTG
                     start"
                     /codon_start=1
                     /transl_table=11
                     /product="orf, hypothetical protein"
                     /protein_id="AAC73208.1"
                     /db_xref="GI:1786286"
                     /translation="MLWTAGFNDKICALNTFEYDRDGNNVSGILTRWRQFGKRYFWPH
                     LLLGMVAASLGLPALSNAAEPNAPAKATTRNHEPSAKVNFGQLALLEANTRRPNSNYS
                     VDYWHQHAIRTVIRHLSFAMAPQTLPVAEESLPLQAQHLALLDTLSALLTQEGTPSEK
                     GYRIDYAHFTPQAKFSTPVWISQAQGIRAGPQRLT"
     promoter        2875. .2907
                     /gene="yacA"
                     /note="factor Sigma70; predicted +1 start at 108158"
     gene            3035. .5740
                     /gene="secA"
                     /note="b0098"
     CDS             3035. .5740
                     /gene="secA"
                     /function="transport; Protein, peptide secretion"
                     /note="o901; 99 pct identical to SECA_ECOLI SW: P10408"
                     /codon_start=1
                     /transl_table=11
                     /product="preprotein translocase; secretion protein"
                     /protein_id="AAC73209.1"
                     /db_xref="GI:1786287"
                     /translation="MLIKLLTKVFGSRNDRTLRRMRKVVNIINAMEPEMEKLSDEELK
                     GKTAEFRARLEKGEVLENLIPEAFAVVREASKRVFGMRHFDVQLLGGMVLNERCIAEM
                     RTGEGKTLTATLPAYLNALTGKGVHVVTVNDYLAQRDAENNRPLFEFLGLTVGINLPG
                     MPAPAKREAYAADITYGTNNEYGFDYLRDNMAFSPEERVQRKLHYALVDEVDSILIDE
                     ARTPLIISGPAEDSSEMYKRVNKIIPHLIRQEKEDSETFQGEGHFSVDEKSRQVNLTE
                     RGLVLIEELLVKEGIMDEGESLYSPANIMLMHHVTAALRAHALFTRDVDYIVKDGEVI
```

```
                     IVDEHTGRTMQGRRWSDGLHQAVEAKEGVQIQNENQTLASITFQNYFRLYEKLAGMTG
                     TADTEAFEFSSIYKLDTVVVPTNRPMIRKDLPDLVYMTEAEKIQAIIEDIKERTAKGQ
                     PVLVGTISIEKSELVSNELTKAGIKHNVLNAKFHANEAAIVAQAGYPAAVTIATNMAG
                     RGTDIVLGGSWQAEVAALENPTAEQIEKIKADWQVRHDAVLEAGGLHIIGTERHESRR
                     IDNQLRGRSGRQGDAGSSRFYLSMEDALMRIFASDRVSGMMRKLGMKPGEAIEHPWVT
                     KAIANAQRKVESRNFDIRKQLLEYDDVANDQRRAIYSQRNELLDVSDVSETINSIRED
                     VFKATIDAYIPPQSLEEMWDIPGLQERLKNDFDLDLPIAEWLDKEPELHEETLRERIL
                     AQSIEVYQRKEEVVGAEMMRHFEKGVMLQTLDSLWKEHLAAMDYLRQGIHLRGYAQKD
                     PKQEYKRESFSMFAAMLESLKYEVISTLSKVQVRMPEEVEELEQQRRMEAERLAQMQQ
                     LSHQDDDSAAAAALAAQTGERKVGRNDPCPCGSGKKYKQCHGRLQ"
     promoter        5756. .5785
                     /note="factor Sigma70; predicted +1 start at 111036"
     gene            5800. .6189
                     /gene="mutT"
                     /note="b0099"
     CDS             5800. .6189
                     /gene="mutT"
                     /EC_number="3.6.1.-"
                     /function="enzyme; 2'-Deoxyribonucleotide metabolism"
                     /note="o129; 100 pct identical to MUTT_ECOLI SW: P08337"
                     /codon_start=1
                     /transl_table=11
                     /product="7,8-dihydro-8-oxoguanine-triphosphatase, prefers
                     dGTP, causes AT-GC transversions"
                     /protein_id="AAC73210.1"
                     /db_xref="GI:1786288"
                     /translation="MKKLQIAVGIIRNENNEIFITRRAADAHMANKLEFPGGKIEMGE
                     TPEQAVVRELQEEVGITPQHFSLFEKLEYEFPDRHITLWFWLVERWEGEPWGKEGQPG
                     EWMSLVGLNADDFPPANEPVIAKLKRL"
     repeat_region   6186. .6385
                     /note="REP (repetitive extragenic palindromic) element;
                     contains 4 REP sequences"
     gene            complement(6320. .6454)
                     /gene="b0100"
     CDS             complement(6320. .6454)
                     /gene="b0100"
                     /function="orf; Unknown"
                     /note="f44; 61 pct identical (1 gap) to 21 residues of an
                     approx. 216 aa protein KTHY_HUMAN SW: P23919"
                     /codon_start=1
                     /transl_table=11
                     /product="orf, hypothetical protein"
                     /protein_id="AAC73211.1"
                     /db_xref="GI:1786289"

  Query Match           100.0%;  Score 20;  DB 1;  Length 12434;
  Best Local Similarity  100.0%;  Pred. No. 62;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GATTAGCATAATAAAATCTC 20
           ||||||||||||||||||||
Db    3043 GATTAGCATAATAAAATCTC 3024


RESULT 8
AE005186/c
LOCUS       AE005186               12518 bp    DNA     linear   BCT 21-MAR-2001
DEFINITION  Escherichia coli O157:H7 EDL933 genome, contig 1 of 3, section 10
            of 155.
ACCESSION   AE005186 AE005174
VERSION     AE005186.1  GI:12512796
KEYWORDS    .
SOURCE      Escherichia coli O157:H7 EDL933.
  ORGANISM  Escherichia coli O157:H7 EDL933
            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
            Escherichia.
REFERENCE   1  (bases 1 to 12518)
  AUTHORS   Perna,N.T., Plunkett,G. III, Burland,V., Mau,B., Glasner,J.D.,
            Rose,D.J., Mayhew,G.F., Evans,P.S., Gregor,J., Kirkpatrick,H.A.,
            Posfai,G., Hackett,J., Klink,S., Boutin,A., Shao,Y., Miller,L.,
            Grotbeck,E.J., Davis,N.W., Lim,A., Dimalanta,E., Potamousis,K.,
            Apodaca,J., Anantharaman,T.S., Lin,J., Yen,G., Schwartz,D.C.,
            Welch,R.A. and Blattner,F.R.
  TITLE     Genome sequence of enterohaemorrhagic Escherichia coli O157:H7
  JOURNAL   Nature 409 (6819), 529-533 (2001)
  MEDLINE   21074935
   PUBMED   11206551
REFERENCE   2  (bases 1 to 12518)
  AUTHORS   Perna,N.T., Plunkett,G. III, Burland,V., Mau,B., Glasner,J.D.,
            Rose,D.J., Mayhew,G.F., Evans,P.S., Gregor,J., Kirkpatrick,H.A.,
            Posfai,G., Hackett,J., Klink,S., Boutin,A., Shao,Y., Miller,L.,
            Grotbeck,E.J., Davis,N.W., Lim,A., Dimalanta,E., Potamousis,K.,
            Apodaca,J., Anantharaman,T.S., Lin,J., Yen,G., Schwartz,D.C.,
            Welch,R.A. and Blattner,F.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-OCT-2000) Laboratory of Genetics, University of
            Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
FEATURES             Location/Qualifiers
     source          1. .12518
                     /organism="Escherichia coli O157:H7 EDL933"
                     /strain="EDL933"
                     /serotype="O157:H7"
                     /db_xref="taxon:155864"
                     /note="enterohemorrhagic"
     gene            38. .1513
                     /gene="murC"
                     /note="Z0101"
     CDS             38. .1513
                     /gene="murC"
                     /function="enzyme; Cell envelop: Murein sacculus,
                     peptidoglycan"
                     /note="Residues 1 to 491 of 491 are 99.79 pct identical to
                     residues 1 to 491 of 491 from Escherichia coli K-12 Strain
                     MG1655: B0091"
                     /codon_start=1
                     /transl_table=11
                     /product="L-alanine adding enzyme,
                     UDP-N-acetyl-muramate:alanine ligase"
                     /protein_id="AAG54395.1"
                     /db_xref="GI:12512797"
                     /translation="MNTQQLAKLRSIVPEMRRVRHIHFVGIGGAGMGGIAEVLANEGY
                     QISGSDLAPNPVTQQLMNLGATIYFNHRPENVRDASVVVVSSAISADNPEIVAAHEAR
                     IPVIRRAEMLAELMRFRHGIAIAGTHGKTTTTAMVSSIYAEAGLDPTFVNGGLVKAAG
                     VHARLGHGRYLIAEADESDASFLHLQPMVAIVTNIEADHMDTYQGDFENLKQTFINFL
                     HNLPFYGRAVMCVDDPVIRELLPRVGRQTTTYGFSEDADVRVEDYQQIGPQGHFTLLR
                     QDKEPMRVTLNAPGRHNALNAAAAVAVATEEGIDDEAILRALESFQGTGRRFDFLGEF
                     PLEPVNGKSGTAMLVDDYGHHPTEVDATIKAARAGWPDKNLVMLFQPHRFTRTRDLYD
                     DFANVLTQVDTLLMLEVYPAGEAPIPGADSRSLCRTIRGRGKIDPILVPDPAQVAEML
                     APVLTGNDLILVQGAGNIGKIARSLAEIKLKPQTPEEEQHD"
     gene            1506. .2426
                     /gene="ddlB"
                     /note="Z0102"
     CDS             1506. .2426
                     /gene="ddlB"
                     /function="enzyme; Cell envelop: Murein sacculus,
                     peptidoglycan"
                     /note="Residues 1 to 306 of 306 are 99.34 pct identical to
                     residues 1 to 306 of 306 from Escherichia coli K-12 Strain
                     MG1655: B0092"
                     /codon_start=1
                     /transl_table=11
                     /product="D-alanine-D-alanine ligase B, affects cell
                     division"
                     /protein_id="AAG54396.1"
                     /db_xref="GI:12512798"
                     /translation="MTDKIAVLLGGTSAEREVSLNSGAAVLAGLREGGIDAYPVDPKE
                     VDVTQLKSMGFQKVFIALHGRGGEDGTLQGMLELMGLPYTGSGVMASALSMDKLRSKL
                     LWQGAGLPVAPWVALTRVEFEKGLSDKQLAEISALGLPVIVKPSREGSSVGMSKVVAE
                     NALQDALRLAFQHDEEVLIEKWLSGPEFTVAILGEEILPSVRIQPSGTFYDYEAKYLS
                     DETQYFCPAGLEASQEANLQALVLKAWTTLGCKGWGRIDVMLDSDGQFYLLEANTSPG
                     MTSHSLVPMAARQAGMSFSQLVVRILELAD"
     gene            2428. .3258
                     /gene="ftsQ"
                     /note="Z0103"
     CDS             2428. .3258
                     /gene="ftsQ"
```

```
                     /function="phenotype; Cell division"
                     /note="Residues 1 to 276 of 276 are 99.63 pct identical to
                     residues 1 to 276 of 276 from Escherichia coli K-12 Strain
                     MG1655: B0093"
                     /codon_start=1
                     /transl_table=11
                     /product="cell division protein; ingrowth of wall at
                     septum"
                     /protein_id="AAG54397.1"
                     /db_xref="GI:12512799"
                     /translation="MSQAALNTRNSEEEVSSRRNNGTRLAGILFLLTVLTTVLVSGWV
                     VLGWMEDAQRLPLSKLVLTGERHYTRNDDIRQSILALGEPGTFMTQDVNIIQTQIEQR
                     LPWIKQVSVRKQWPDELKIHLVEYVPIARWNDQHMVDAEGNTFSVPPDRTSKQVLPML
                     YGPEGSANEVLQGYREMGQMLAKDRFTLKEAAMTARRSWQLTLNNDIKLNLGRGDTMK
                     RLARFVELYPVLQQQAQTDGKRISYVDLRYDSGAAVGWAPLPPEESTQQQNQAQAEQQ
                     "
gene                 3255. .4517
                     /gene="ftsA"
                     /note="Z0104"
CDS                  3255. .4517
                     /gene="ftsA"
                     /function="phenotype; Cell division"
                     /note="Residues 1 to 420 of 420 are 100.00 pct identical
                     to residues 1 to 420 of 420 from Escherichia coli K-12
                     Strain MG1655: B0094"
                     /codon_start=1
                     /transl_table=11
                     /product="ATP-binding cell division protein, septation
                     process, complexes with FtsZ, associated with junctions of
                     inner and outer membranes"
                     /protein_id="AAG54398.1"
                     /db_xref="GI:12512800"
                     /translation="MIKATDRKLVVGLEIGTAKVAALVGEVLPDGMVNIIGVGSCPSR
                     GMDKGGVNDLESVVKCVQRAIDQAELMADCQISSVYLALSGKHISCQNEIGMVPISEE
                     EVTQEDVENVVHTAKSVRVRDEHRVLHVIPQEYAIDYQEGIKNPVGLSGVRMQAKVHL
                     ITCHNDMAKNIVKAVERCGLKVDQLIFAGLASSYSVLTEDERELGVCVVDIGGGTMDI
                     AVYTGGALRHTKVIPYAGNVVTSDIAYAFGTPPSDAEAIKVRHGCALGSIVGKDESVE
                     VPSVGGRPPRSLQRQTLAEVIEPRYTELLNLVNEEILQLQEKLRQQGVKHHLAAGIVL
                     TGGAAQIEGLAACAQRVFHTQVRIGAPLNITGLTDYAQEPYYSTAVGLLHYGKESHLN
                     GEAEVEKRVTASVGSWIKRLNSWLRKEF"
gene                 4578. .5729
                     /gene="ftsZ"
                     /note="Z0105"
CDS                  4578. .5729
                     /gene="ftsZ"
                     /function="enzyme; Cell division"
                     /note="Residues 1 to 383 of 383 are 100.00 pct identical
                     to residues 1 to 383 of 383 from Escherichia coli K-12
                     Strain MG1655: B0095"
                     /codon_start=1
                     /transl_table=11
                     /product="cell division; forms circumferential ring;
                     tubulin-like GTP-binding protein and GTPase"
                     /protein_id="AAG54399.1"
                     /db_xref="GI:12512801"
                     /translation="MFEPMELTNDAVIKVIGVGGGGGNAVEHMVRERIEGVEFFAVNT
                     DAQALRKTAVGQTIQIGSGITKGLGAGANPEVGRNAADEDRDALRAALEGADMVFIAA
                     GMGGGTGTGAAPVVAEVAKDLGILTVAVVTKPFNFEGKKRMAFAEQGITELSKHVDSL
                     ITIPNDKLLKVLGRGISLLDAFGAANDVLKGAVQGIAELITRPGLMNVDFADVRTVMS
                     EMGYAMMGSGVASGEDRAEEAAEMAISSPLLEDIDLSGARGVLVNITAGFDLRLDEFE
                     TVGNTIRAFASDNATVVIGTSLDPDMNDELRVTVVATGIGMDKRPEITLVTNKQVQQP
                     VMDRYQQHGMAPLTQEQKPVAKVVNDNAPQTAKEPDYLDIPAFLRKQAD"
gene                 5830. .6747
                     /gene="lpxC"
                     /note="Z0106"
CDS                  5830. .6747
                     /gene="lpxC"
                     /function="enzyme; Cell exterior constituents: Surface
                     polysaccharides and antigens"
                     /note="Residues 1 to 305 of 305 are 100.00 pct identical
                     to residues 1 to 305 of 305 from Escherichia coli K-12
                     Strain MG1655: B0096"
                     /codon_start=1
                     /transl_table=11
                     /product="UDP-3-O-acyl N-acetylglucosamine deacetylase;
                     lipid A biosynthesis"
                     /protein_id="AAG54400.1"
                     /db_xref="GI:12512802"
                     /translation="MIKQRTLKRIVQATGVGLHTGKKVTLTLRPAPANTGVIYRRTDL
                     NPPVDFPADAKSVRDTMLCTCLVNEHDVRISTVEHLNAALAGLGIDNIVIEVNAPEIP
                     IMDGSAAPFVYLLLDAGIDELNCAKKFVRIKETVRVEDGDKWAEFKPYNGFSLDFTID
                     FNHPAIDSSNQRYAMNFSADAFMRQISRARTFGFMRDIEYLQSRGLCLGGSFDCAIVV
                     DDYRVLNEDGLRFEDEFVRHKMLDAIGDLFMCGHNIIGAFTAYKSGHALNNKLLQAVL
                     AKQEAWEYVTFQDDAELPLAFKAPSAVLA"
gene                 6903. .7490
                     /gene="yacA"
                     /note="Z0107"
CDS                  6903. .7490
                     /gene="yacA"
                     /function="orf; Unknown function"
                     /note="Residues 1 to 195 of 195 are 98.97 pct identical to
                     residues 1 to 195 of 195 from Escherichia coli K-12 Strain
                     MG1655: B0097"
                     /codon_start=1
                     /transl_table=11
                     /product="orf, hypothetical protein"
                     /protein_id="AAG54401.1"
                     /db_xref="GI:12512803"
                     /translation="MLWTSGFNDKICALNTFEYDRDGNNVSGILTRWRQFGKRYFWPH
                     LLLGMVAASLGLPALSNAAEPNAPAKATTRNHEPSAKVNFGQLALLEANTRRPNSNYS
                     VDYWHQHAIRTVIRHLSFAMAPQTLPVAEESLPLQAQHLALLDTLSALLTQEGTPSEK
                     GYRIDYAHFTPQAKFSTPVWISQAQGIRAGPQRLS"
gene                 7552. .10257
                     /gene="secA"
                     /note="Z0108"
CDS                  7552. .10257
                     /gene="secA"
                     /function="transport; Transport of large molecules:
                     Protein, peptide secretion"
                     /note="Residues 1 to 901 of 901 are 99.88 pct identical to

  Query Match          100.0%;  Score 20;  DB 1;  Length 12518;
  Best Local Similarity  100.0%;  Pred. No. 62;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GATTAGCATAATAAAATCTC 20
           ||||||||||||||||||||
Db    7560 GATTAGCATAATAAAATCTC 7541


RESULT 9
EC2MIN/c
LOCUS       EC2MIN                28277 bp    DNA     linear   BCT 05-JUL-1999
DEFINITION  E. coli 2 minute region.
ACCESSION   X55034 M10429
VERSION     X55034.1  GI:40841
KEYWORDS    ddl gene; envA gene; ftsA gene; ftsQ gene; ftsW gene; ftsZ gene;
            ilvH gene; ilvI gene; leuA gene; leuO gene; mraY gene; murC gene;
            murD gene; murE gene; murF gene; murG gene; mutT gene; orfA; orfB;
            orfC; orfX; pbpB gene; secA gene; shl gene.
SOURCE      Escherichia coli.
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
            Escherichia.
REFERENCE   1  (bases 1 to 306)
  AUTHORS   Wessler,S.R. and Calvo,J.M.
  TITLE     Control of leu operon expression in Escherichia coli by a
            transcription attenuation mechanism
  JOURNAL   J. Mol. Biol. 149, 579-579 (1981)
  MEDLINE   82078077
REFERENCE   2  (bases 2101 to 4431)
  AUTHORS   Squires,C.H., DeFelice,M., Devereux,J. and Calvo,J.M.
  TITLE     Molecular structure of ilvIH and its evolutionary relationship to
            ilvG in Escherichia coli
  JOURNAL   Nucleic Acids Res. 11, 5299-5313 (1983)
  MEDLINE   83272971
```

REFERENCE 3 (bases 7316 to 10074)
AUTHORS Nakamura,M., Maruyama,I.N., Soma,M., Kato,J.I., Suzuki,H. and Hirota,Y.
TITLE On the process of cellular division in Escherichia coli: Nucleotide sequence of the gene for penicillin-binding protein 3
JOURNAL Mol. Gen. Genet. 191, 1-9 (1983)
MEDLINE 83296957
REFERENCE 4 (bases 19464 to 21952)
AUTHORS Robinson,A.C., Kenan,D.J., Hatfull,G.F., Sullivan,N.F., Spiegelberg,R. and Donachie,W.D.
TITLE DNA sequence and transcriptional organization of essential cell division genes ftsQ and ftsA of Escherichia coli: Evidence for overlapping transcriptional units
JOURNAL J. Bacteriol. 160, 546-555 (1984)
MEDLINE 85054557
REFERENCE 5 (bases 21464 to 23333)
AUTHORS Yi,Q.M. and Lutkenhaus,J.
TITLE The nucleotide sequence of the essential cell-division gene ftsZ of Escherichia coli
JOURNAL Gene 36, 241-247 (1985)
MEDLINE 86083166
REFERENCE 6 (bases 1799 to 2187)
AUTHORS Haughn,G.W., Squires,C.H., DeFelice,M., Largo,C.T. and Calvo,J.M.
TITLE Unusual organization of the ilvIH promoter of Escherichia coli
JOURNAL J. Bacteriol. 163, 186-198 (1985)
MEDLINE 85234358
REFERENCE 7 (bases 268 to 1130)
AUTHORS Haughn,G.W., Wessler,S.R., Gemmill,R.M. and Calvo,J.M.
TITLE High A + T content conserved in DNA sequences upstream of leuABCD in Escherichia coli and Salmonella typhimurium
JOURNAL J. Bacteriol. 166, 1113-1117 (1986)
MEDLINE 86223773
REFERENCE 8 (bases 18619 to 19770)
AUTHORS Robinson,A.C., Kenan,D.J., Sweeney,J. and Donachie,W.D.
TITLE Further Evidence for Overlapping Transcriptional Units in an Escherichia coli Cell Envelope-Cell Division Gene Cluster: DNA Sequence and Transcriptional Organization of the ddl ftsQ Region
JOURNAL J. Bacteriol. 167, 809-817 (1986)
MEDLINE 86304170
REFERENCE 9 (bases 22964 to 25011)
AUTHORS Beall,B. and Lutkenhaus,J.
TITLE Sequence analysis, transcriptional organization, and insertional mutagenesis of the envA gene of Escherichia coli
JOURNAL J. Bacteriol. 169, 5408-5415 (1987)
MEDLINE 88058745
REFERENCE 10 (bases 27412 to 28277)
AUTHORS Akiyama,M., Horiuchi,T. and Sekiguchi,M.
TITLE Molecular cloning and nucleotide sequence of the mutT mutator of Escherichia coli that causes A:T to C:G transversion
JOURNAL Mol. Gen. Genet. 206, 9-16 (1987)
MEDLINE 87201091
REFERENCE 11 (bases 843 to 1812)
AUTHORS Henikoff,S., Haughn,G.W., Calvo,J.M. and Wallace,J.C.
TITLE A large family of bacterial activator proteins
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 85, 6602-6606 (1988)
MEDLINE 88320486
REFERENCE 12 (bases 20513 to 21772)
AUTHORS Robinson,A.C., Begg,K.J. and Donachie,W.D.
TITLE Mapping and characterization of mutants of the Escherichia coli cell division gene, ftsA
JOURNAL Mol. Microbiol. 2, 581-588 (1988)
MEDLINE 89039246
REFERENCE 13 (bases 23989 to 27799)
AUTHORS Schmidt,M., Rollo,E., Grodberg,I. and Oliver,D.
TITLE Nucleotide sequence of secA gene and secA(ts) mutations preventing protein export in Escherichia coli
JOURNAL J. Bacteriol. 170, 3404-3414 (1988)
MEDLINE 88298644
REFERENCE 14 (bases 11142 to 12634)
AUTHORS Parquet,C., Fluoret,B., Menguin-Lecreulx,D. and Van Heijenoort,J.
TITLE Nucleotide sequence of the murF gene encoding the UDP-MurNac-pentapeptide synthetase of Escherichia coli
JOURNAL Nucleic Acids Res. 17, 5379-5379 (1989)
MEDLINE 89345095
REFERENCE 15 (bases 14743 to 16239)
AUTHORS Ikeda,M., Sato,T., Wachi,M., Jung,H.K., Ishino,F., Kobayashi,Y. and Matsuhashi,M.
TITLE Structural similarity among Escherichia coli FtsW and RodA proteins and Bacillus subtilis SpoVE protein, which function in cell division, cell elongation, and spore formation, respectively
JOURNAL J. Bacteriol. 171, 6375-6378 (1989)
MEDLINE 90036736
REFERENCE 16 (bases 1 to 28277)
AUTHORS Tao,J.S. and Ishiguro,E.E.
TITLE Nucleotide sequence of the murE gene of Escherichia coli
JOURNAL Can. J. Microbiol. 35, 1051-1054 (1989)
MEDLINE 90124047
REFERENCE 17 (bases 13392 to 15020)
AUTHORS Mengin-Lecreulx,D. and van Heijenoort,J.
TITLE Nucleotide sequence of the murD gene encoding the UDP-MurNAc-L-Ala-D-Glu synthetase of Escherichia coli
JOURNAL Nucleic Acids Res. 18 (1), 183 (1990)
MEDLINE 90174916
REFERENCE 18 (bases 12423 to 15030)
AUTHORS Ikeda,M., Wachi,M., Ishino,F. and Matsuhashi,M.
TITLE Nucleotide sequence involving murD and an open reading frame orf-Y spacing murF and ftsW in Escherichia coli
JOURNAL Nucleic Acids Res. 18, 1058-1058 (1990)
MEDLINE 90192099
REFERENCE 19 (bases 6088 to 7587)
AUTHORS Gomez,M.J., Fluoret,B., Van Heijenoort,J. and Ayala,J.A.
TITLE Nucleotide sequence of the regulatory region of pbpB gene of Escherichia coli
JOURNAL Nucleic Acids Res. 18, 2813-2813 (1990)
MEDLINE 90251464
REFERENCE 20 (bases 4274 to 6093)
AUTHORS Leclerc,G., Noel,G. and Drapeau,G.
TITLE Molecular cloning, nucleotide sequence and expression of shl, a new gene in the 2-minute region of the genetic map of Escherichia coli
JOURNAL J. Bacteriol. 172, 4696-4700 (1990)
MEDLINE 90330585
REFERENCE 21 (bases 16094 to 18886)
AUTHORS Ikeda,M., Wachi,M., Jung,H.K., Ishino,F. and Matsuhashi,M.
TITLE Nucleotide sequence involving murG and murC in the mra gene cluster region of Escherichia coli
JOURNAL Nucleic Acids Res. 18, 4014-4014 (1990)
MEDLINE 90326550
REFERENCE 22 (bases 16094 to 17806)
AUTHORS Mengin-Lecreulx,D., Texier,L. and van Heijenoort,J.
TITLE Nucleotide sequence of the cell-envelope murG gene of Escherichia coli
JOURNAL Nucleic Acids Res. 18 (9), 2810 (1990)
MEDLINE 90251461
REFERENCE 23 (bases 1 to 28277)
AUTHORS Michaud,C., Parquet,C., Flouret,B., Blanot,D. and van Heijenoort,J.
TITLE Revised interpretation of the sequence containing the murE gene encoding the UDP-N-acetylmuramyl-tripeptide synthetase of Escherichia coli
JOURNAL Biochem. J. 269 (1), 277-278 (1990)
MEDLINE 90328986
REFERENCE 24 (bases 1 to 28277)
AUTHORS Wang,Q. and Calvo,J.M.
TITLE Lrp a global regulatory protein of escherichia coli binds Co-operatively to multiple sites and activates transcrioption of ilvIH
JOURNAL J. Mol. Biol. 229, 306-318 (1993)
MEDLINE 93156044
REFERENCE 25 (bases 1 to 28277)
AUTHORS Wang,Q. and Calvo,J.M.
TITLE Lrp a major regulatory protein in Escherichia coli bends DNA and can organize the assembly of a higher-order nucleoprotein structure
JOURNAL EMBO J. 12, 2495-2501 (1993)
MEDLINE 93285120
REFERENCE 26 (bases 4274 to 6093)
AUTHORS Jahreis,K., Postma,P.W. and Lengeler,J.W.
TITLE Nucleotide sequence of the ilv H-fruR gene of Escherichia coli K-12

```
            and Salmonella typhimurium LT2
  JOURNAL   Unpublished
REFERENCE   27 (bases 1 to 28277)
  AUTHORS   Ayala,J.A.
  TITLE     Regulation of transcription at the 2-minute region of the genetic
            map of Escherichia coli
  JOURNAL   Unpublished
REFERENCE   28 (bases 1 to 28277)
  AUTHORS   Ayala,J.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JAN-1991) Ayala J.A., Instituto de Biologia
            Molecular, Centro de Biologia Molecular, Universidad Autonoma,
            Canto-Blanco 28049, Madrid, Spain
COMMENT     This entry comprises a merged sequence of 28Kb of which a portion
            is the submittor's original work.
FEATURES             Location/Qualifiers
     source          1. .28277
                     /organism="Escherichia coli"
                     /strain="K12 W3110"
                     /db_xref="taxon:562"
                     /map="1.77 to 2.38 mins"
                     /clone="pLC26-6, pLC4-14, pLC6-7"
     mRNA            complement(<1. .269)
                     /product="leu mRNA"
                     /evidence=experimental
     mat_peptide     complement(<1. .63)
                     /gene="leuA"
                     /product="alpha-isopropylmalate synthase (LeuA)"
     gene            complement(1. .63)
                     /gene="leuA"

  Query Match             100.0%;  Score 20;  DB 1;  Length 28277;
  Best Local Similarity   100.0%;  Pred. No. 52;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GATTAGCATAATAAAATCTC 20
          ||||||||||||||||||||
Db  24817 GATTAGCATAATAAAATCTC 24798


RESULT 10
AX191720/c
LOCUS       AX191720               28277 bp    DNA     linear   PAT 15-AUG-2001
DEFINITION  Sequence 2 from Patent WO0149775.
ACCESSION   AX191720
VERSION     AX191720.1  GI:15209889
KEYWORDS    .
SOURCE      Escherichia coli.
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
            Escherichia.
REFERENCE   1  (bases 1 to 28277)
  AUTHORS   Iversen,P.L.
  TITLE     Antisense antibacterial cell division composition and method
  JOURNAL   Patent: WO 0149775-A 2 12-JUL-2001;
            Avi Biopharma, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .28277
                     /organism="Escherichia coli"
                     /db_xref="taxon:562"
BASE COUNT     6714 a   6725 c   7993 g   6845 t
ORIGIN

  Query Match             100.0%;  Score 20;  DB 6;  Length 28277;
  Best Local Similarity   100.0%;  Pred. No. 52;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 GATTAGCATAATAAAATCTC 20
          ||||||||||||||||||||
Db  24817 GATTAGCATAATAAAATCTC 24798


RESULT 11
ECO110K/c
LOCUS       ECO110K               111401 bp    DNA     linear   BCT 27-JUN-1997
DEFINITION  E.coli K12 genome, 0-2.4min. region.
ACCESSION   D10483 J01597 J01683 J01706 K01298 K01990 M10420 M10611 M12544
            V00259 X04711 X54847 X54945 X55034 X56742
VERSION     D10483.1  GI:216434
KEYWORDS    Gef; DnaJ.
SOURCE      Escherichia coli (strain:K-12) DNA.
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
            Escherichia.
REFERENCE   1  (bases 1 to 111401)
  AUTHORS   Mori,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-FEB-1992) Hirotada Mori, Institute for Virus
            Research, Kyoto University, Genetics and Molecular Biology; 53
            Shogoin Kawara-Machi, Sakyo-ku, Kyoto 606, Japan
            (E-mail:e52985@sakura.kudpc.kyoto-u.ac.jp, Tel:075-751-4042,
            Fax:075-761-5626)
REFERENCE   2  (sites)
  AUTHORS   Smith,B.R. and Schleif,R.
  TITLE     Nucleotide sequence of the L-arabinose regulatory region of
            Escherichia coli K12
  JOURNAL   The Journal of biological chemistry. 253 (19), 6931-6933 (1978)
  MEDLINE   79005683
   PUBMED   357433
REFERENCE   3  (sites)
  AUTHORS   Ohtsubo,H. and Ohtsubo,E.
  TITLE     Nucleotide sequence of an insertion element, IS1
  JOURNAL   Proceedings of the National Academy of Sciences of the United
            States of America. 75 (2), 615-619 (1978)
  MEDLINE   78137003
   PUBMED   273224
REFERENCE   4  (sites)
  AUTHORS   Greenfield,L., Boone,T. and Wilcox,G.
  TITLE     DNA sequence of the araBAD promoter in Escherichia coli B/r
  JOURNAL   Proceedings of the National Academy of Sciences of the United
            States of America. 75 (10), 4724-4728 (1978)
  MEDLINE   79116194
   PUBMED   368797
REFERENCE   5  (sites)
  AUTHORS   Johnsrud,L.
  TITLE     DNA sequence of the transposable element IS1
  JOURNAL   Molecular & general genetics : MGG. 169 (2), 213-218 (1979)
  MEDLINE   79177885
   PUBMED   375010
REFERENCE   6  (sites)
  AUTHORS   Smith,D.R. and Calvo,J.M.
  TITLE     Nucleotide sequence of the E coli gene coding for dihydrofolate
            reductase
  JOURNAL   Nucleic acids research. 8 (10), 2255-2274 (1980)
  MEDLINE   81053692
   PUBMED   6159575
REFERENCE   7  (sites)
  AUTHORS   Miyada,C.G., Horwitz,A.H., Cass,L.G., Timko,J. and Wilcox,G.
  TITLE     DNA sequence of the araC regulatory gene from Escherichia coli B/r
  JOURNAL   Nucleic acids research. 8 (22), 5267-5274 (1980)
  MEDLINE   81124262
   PUBMED   7008027
REFERENCE   8  (sites)
  AUTHORS   Ogden,S., Haggerty,D., Stoner,C.M., Kolodrubetz,D. and Schleif,R.
  TITLE     The Escherichia coli L-arabinose operon: binding sites of the
            regulatory proteins and a mechanism of positive and negative
            regulation
  JOURNAL   Proceedings of the National Academy of Sciences of the United
            States of America. 77 (6), 3346-3350 (1980)
  MEDLINE   81013881
   PUBMED   6251457
REFERENCE   9  (sites)
  AUTHORS   Katinka,M., Cossart,P., Sibilli,L., Saint-Girons,I.,
            Chalvignac,M.A., Le Bras,G., Cohen,G.N. and Yaniv,M.
  TITLE     Nucleotide sequence of the thrA gene of Escherichia coli
  JOURNAL   Proceedings of the National Academy of Sciences of the United
```

States of America. 77 (10), 5730-5733 (1980)
MEDLINE 81077247
PUBMED 7003595
REFERENCE 10 (sites)
AUTHORS Mackie,G.A.
TITLE Nucleotide sequence of the gene for ribosomal protein S20 and its flanking regions
JOURNAL The Journal of biological chemistry. 256 (15), 8177-8182 (1981)
MEDLINE 81264207
PUBMED 6267039
REFERENCE 11 (sites)
AUTHORS Cossart,P., Katinka,M. and Yaniv,M.
TITLE Nucleotide sequence of the thrB gene of E. coli, and its two adjacent regions; the thrAB and thrBC junctions
JOURNAL Nucleic acids research. 9 (2), 339-347 (1981)
MEDLINE 81150470
PUBMED 6259626
REFERENCE 12 (sites)
AUTHORS Lee,N.L., Gielow,W.O. and Wallace,R.G.
TITLE Mechanism of araC autoregulation and the domains of two overlapping promoters, Pc and PBAD, in the L-arabinose regulatory region of Escherichia coli
JOURNAL Proceedings of the National Academy of Sciences of the United States of America. 78 (2), 752-756 (1981)
MEDLINE 81199399
PUBMED 6262769
REFERENCE 13 (sites)
AUTHORS Stoner,C.M. and Schleif,R.
TITLE Is the amino acid but not the nucleotide sequence of the Escherichia coli araC gene conserved?
JOURNAL Journal of molecular biology. 154 (4), 649-652 (1982)
MEDLINE 82216830
PUBMED 6283093
REFERENCE 14 (sites)
AUTHORS Gilson,E., Nikaido,H. and Hofnung,M.
TITLE Sequence of the malK gene in E.coli K12
JOURNAL Nucleic acids research. 10 (22), 7449-7458 (1982)
MEDLINE 83116968
PUBMED 6296778
REFERENCE 15 (sites)
AUTHORS Parsot,C., Cossart,P., Saint-Girons,I. and Cohen,G.N.
TITLE Nucleotide sequence of thrC and of the transcription termination region of the threonine operon in Escherichia coli K12
JOURNAL Nucleic acids research. 11 (21), 7331-7345 (1983)
MEDLINE 84069770
PUBMED 6316258
REFERENCE 16 (sites)
AUTHORS Bouvier,J., Richaud,C., Richaud,F., Patte,J.C. and Stragier,P.
TITLE Nucleotide sequence and expression of the Escherichia coli dapB gene
JOURNAL The Journal of biological chemistry. 259 (23), 14829-14834 (1984)
MEDLINE 85054974
PUBMED 6094578
REFERENCE 17 (sites)
AUTHORS Bardwell,J.C. and Craig,E.A.
TITLE Major heat shock gene of Drosophila and the Escherichia coli heat-inducible dnaK gene are homologous
JOURNAL Proceedings of the National Academy of Sciences of the United States of America. 81 (3), 848-852 (1984)
MEDLINE 84144800
PUBMED 6322174
REFERENCE 18 (sites)
AUTHORS Innis,M.A., Tokunaga,M., Williams,M.E., Loranger,J.M., Chang,S.Y., Chang,S. and Wu,H.C.
TITLE Nucleotide sequence of the Escherichia coli prolipoprotein signal peptidase (lsp) gene
JOURNAL Proceedings of the National Academy of Sciences of the United States of America. 81 (12), 3708-3712 (1984)
MEDLINE 84222028
PUBMED 6374664
REFERENCE 19 (sites)
AUTHORS Bouvier,J., Patte,J.C. and Stragier,P.
TITLE Multiple regulatory signals in the control region of the Escherichia coli carAB operon
JOURNAL Proceedings of the National Academy of Sciences of the United States of America. 81 (13), 4139-4143 (1984)
MEDLINE 84248073
PUBMED 6377309
REFERENCE 20 (sites)
AUTHORS Chong,P., Hui,I., Loo,T. and Gillam,S.
TITLE Structural analysis of a new GC-specific insertion element IS186
JOURNAL FEBS letters. 192 (1), 47-52 (1985)
MEDLINE 86030702
PUBMED 2996940
REFERENCE 21 (sites)
AUTHORS Kamio,Y., Lin,C.K., Regue,M. and Wu,H.C.
TITLE Characterization of the ileS-lsp operon in Escherichia coli. Identification of an open reading frame upstream of the ileS gene and potential promoter(s) for the ileS-lsp operon
JOURNAL The Journal of biological chemistry. 260 (9), 5616-5620 (1985)
MEDLINE 85182715
PUBMED 2985604
REFERENCE 22 (sites)
AUTHORS Friedberg,D., Rosenthal,E.R., Jones,J.W. and Calvo,J.M.
TITLE Characterization of the 3' end of the leucine operon of Salmonella typhimurium
JOURNAL Molecular & general genetics : MGG. 199 (3), 486-494 (1985)
MEDLINE 85295470
PUBMED 2993799
REFERENCE 23 (sites)
AUTHORS Cowing,D.W., Bardwell,J.C., Craig,E.A., Woolford,C., Hendrix,R.W. and Gross,C.A.
TITLE Consensus sequence for Escherichia coli heat shock gene promoters
JOURNAL Proceedings of the National Academy of Sciences of the United States of America. 82 (9), 2679-2683 (1985)
MEDLINE 85190560
PUBMED 3887408
REFERENCE 24 (sites)
AUTHORS Sekiguchi,T., Ortega-Cesena,J., Nosoh,Y., Ohashi,S., Tsuda,K. and Kanaya,S.
TITLE DNA and amino-acid sequences of 3-isopropylmalate dehydrogenase of Bacillus coagulans. Comparison with the enzymes of Saccharomyces cerevisiae and Thermus thermophilus
JOURNAL Biochim. Biophys. Acta 867, 36-44 (1986)
REFERENCE 25 (sites)
AUTHORS Lee,N., Gielow,W., Martin,R., Hamilton,E. and Fowler,A.
TITLE The organization of the araBAD operon of Escherichia coli
JOURNAL Gene. 47 (2-3), 231-244 (1986)
MEDLINE 87163495
PUBMED 3549454
REFERENCE 26 (sites)
AUTHORS Ohki,M., Tamura,F., Nishimura,S. and Uchida,H.
TITLE Nucleotide sequence of the Escherichia coli dnaJ gene and purification of the gene product
JOURNAL The Journal of biological chemistry. 261 (4), 1778-1781 (1986)
MEDLINE 86111849
PUBMED 3003084
REFERENCE 27 (sites)

Query Match 100.0%; Score 20; DB 1; Length 111401;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATTAGCATAATAAAATCTC 20
||||||||||||||||||||
Db 107941 GATTAGCATAATAAAATCTC 107922

RESULT 12
AC055113
LOCUS AC055113 155976 bp DNA linear HTG 24-AUG-2000
DEFINITION Homo sapiens chromosome 12 clone CTD-2021H9, WORKING DRAFT SEQUENCE, 19 unordered pieces.
ACCESSION AC055113
VERSION AC055113.2 GI:9838029
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

```
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 155976)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 155976)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-APR-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Aug 17, 2000 this sequence version replaced gi:7579846.

            ---------- Genome Center ----------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            ---------- Project Information ----------
            Center project name: H_MS2021H09
            ---------- Summary Statistics ----------
            Sequencing vector: M13; 100%
            Sequencing vector: plasmid; 0%
            Chemistry: Dye-primer ET; 100% of reads
            Chemistry: Dye-terminator Big Dye; 0% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 143933 bases at least Q40
            Consensus quality: 146482 bases at least Q30
            Consensus quality: 147905 bases at least Q20
            Insert size: 180000; agarose-fp
            Insert size: 154176; sum-of-contigs
            Quality coverage: 5.12 in Q20 bases; agarose-fp
            Quality coverage: 5.01 in Q20 bases; sum-of-contigs
            ----------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 19 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     3520: contig of 3520 bp in length
            *    3521     3620: gap of unknown length
            *    3621     9066: contig of 5446 bp in length
            *    9067     9166: gap of unknown length
            *    9167    13195: contig of 4029 bp in length
            *   13196    13295: gap of unknown length
            *   13296    17854: contig of 4559 bp in length
            *   17855    17954: gap of unknown length
            *   17955    26904: contig of 8950 bp in length
            *   26905    27004: gap of unknown length
            *   27005    35606: contig of 8602 bp in length
            *   35607    35706: gap of unknown length
            *   35707    43003: contig of 7297 bp in length
            *   43004    43103: gap of unknown length
            *   43104    52582: contig of 9479 bp in length
            *   52583    52682: gap of unknown length
            *   52683    62116: contig of 9434 bp in length
            *   62117    62216: gap of unknown length
            *   62217    75369: contig of 13153 bp in length
            *   75370    75469: gap of unknown length
            *   75470    88516: contig of 13047 bp in length
            *   88517    88616: gap of unknown length
            *   88617   104061: contig of 15445 bp in length
            *  104062   104161: gap of unknown length
            *  104162   119796: contig of 15635 bp in length
            *  119797   119896: gap of unknown length
            *  119897   141280: contig of 21384 bp in length
            *  141281   141380: gap of unknown length
            *  141381   142833: contig of 1453 bp in length
            *  142834   142933: gap of unknown length
            *  142934   145082: contig of 2149 bp in length
            *  145083   145182: gap of unknown length
            *  145183   148566: contig of 3384 bp in length
            *  148567   148666: gap of unknown length
            *  148667   152318: contig of 3652 bp in length
            *  152319   152418: gap of unknown length
            *  152419   155976: contig of 3558 bp in length.
FEATURES             Location/Qualifiers
     source          1. .155976
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="12"
                     /clone="CTD-2021H9"
     misc_feature    1. .3520
                     /note="assembly_name:Contig10"
     misc_feature    3621. .9066
                     /note="assembly_name:Contig11"
     misc_feature    9167. .13195
                     /note="assembly_name:Contig12"
     misc_feature    13296. .17854
                     /note="assembly_name:Contig13"
     misc_feature    17955. .26904
                     /note="assembly_name:Contig14"
     misc_feature    27005. .35606
                     /note="assembly_name:Contig15"
     misc_feature    35707. .43003
                     /note="assembly_name:Contig16"
     misc_feature    43104. .52582
                     /note="assembly_name:Contig17"
     misc_feature    52683. .62116
                     /note="assembly_name:Contig18"
     misc_feature    62217. .75369
                     /note="assembly_name:Contig19"
     misc_feature    75470. .88516
                     /note="assembly_name:Contig20"
     misc_feature    88617. .104061
                     /note="assembly_name:Contig21"
     misc_feature    104162. .119796
                     /note="assembly_name:Contig22"
     misc_feature    119897. .141280
                     /note="assembly_name:Contig23"
     misc_feature    141381. .142833
                     /note="assembly_name:Contig5"
     misc_feature    142934. .145082
                     /note="assembly_name:Contig6"
     misc_feature    145183. .148566
                     /note="assembly_name:Contig7"
     misc_feature    148667. .152318
                     /note="assembly_name:Contig8"
     misc_feature    152419. .155976
                     /note="assembly_name:Contig9"
BASE COUNT    48607 a  29348 c  28870 g  47349 t   1802 others
ORIGIN

  Query Match             95.0%;  Score 19;  DB 2;  Length 155976;
  Best Local Similarity   100.0%;  Pred. No. 98;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 ATTAGCATAATAAAATCTC 20
          |||||||||||||||||||
Db  94819 ATTAGCATAATAAAATCTC 94837


RESULT 13
AC084877/c
LOCUS       AC084877              179212 bp    DNA     linear   PRI 05-JUN-2001
DEFINITION  Homo sapiens 12q BAC RP11-158O11 (Roswell Park Cancer Institute
            Human BAC Library) complete sequence.
ACCESSION   AC084877
VERSION     AC084877.18  GI:14277186
KEYWORDS    HTG.
SOURCE      human.
```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 179212)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Marondel,I., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M., Neal,D., Nelson,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shim,C., Shooshtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Kucherlapati,R., Weinstock,G. and Gibbs,R.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 179212)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (28-NOV-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 179212)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (01-JUN-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE 4 (bases 1 to 179212)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (05-JUN-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT On Jun 1, 2001 this sequence version replaced gi:14150309.
INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE:Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

QUALSTAT-REPORT----------

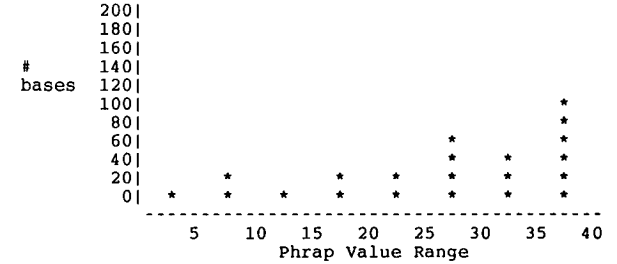
--- Summary Statistics ---

Contig length:	131446
Phrap values in estimate:	131252
Average error rate (BCM-Phrap estimate):	5.08653e-05
Fraction of Phrap values less than 40 :	0.00239996
Number of consensus changing edits:	6
Number of N's in consensus :	0

--- Consensus changing edits ---

Position	Original+Context	Edited+Context
82958	tttcctttgt(t)aaagatctct	tttcctttgt(a)aaanatnnct
82962	ctttgttaaa(g)atctcttcag	ctttgtaaaa(n)atnncttcan
82965	tgttaaagat(c)tcttcagagg	tgtaaaanat(n)ncttcanagg
82966	gttaaagatc(t)cttcagaggc	gtaaaanatn(n)cttcanaggc
82972	gatctcttca(g)aggctaccaa	natnncttca(n)aggctaccaa
82985	gctaccaagg(t)catacccaga	gctaccaagg(n)catacccaga

--- Distribution of Quality < 40 Bases ---



Version: 1.01 qxfo.

FEATURES Location/Qualifiers

```
     source          1. .179212
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="12q"
                     /clone="RP11-158O11"
     repeat_region   complement(489. .762)
                     /rpt_family="MLT1E1"
     repeat_region   886. .1010
                     /rpt_family="L1MA10"
     repeat_region   1299. .1616
                     /rpt_family="AluY"
     repeat_region   2804. .2898
                     /rpt_family="AT_rich"
     repeat_region   3417. .3449
                     /rpt_family="AT_rich"
     repeat_region   3827. .3900
                     /rpt_family="L2"
     repeat_region   4383. .4419
                     /rpt_family="(CA)n"
     repeat_region   complement(4724. .5301)
                     /rpt_family="MLT1E2"
     repeat_region   6920. .6970
                     /rpt_family="(TATG)n"
     repeat_region   complement(7246. .7548)
                     /rpt_family="AluYa5"
     repeat_region   8006. .8050
                     /rpt_family="(TA)n"
     repeat_region   complement(8134. .8458)
                     /rpt_family="LTR16C"
     repeat_region   complement(11541. .11619)
                     /rpt_family="MIR"
     repeat_region   complement(12111. .12183)
                     /rpt_family="MIR"
     repeat_region   12510. .12530
                     /rpt_family="AT_rich"
     repeat_region   complement(13017. .13761)
                     /rpt_family="L1ME3A"
     repeat_region   13770. .14605
                     /rpt_family="L1M4"
     repeat_region   complement(15257. .15326)
                     /rpt_family="MIR"
     repeat_region   15399. .15576
                     /rpt_family="MIR"
     repeat_region   15588. .16122
                     /rpt_family="MLT1G1"
     repeat_region   complement(16159. .16432)
                     /rpt_family="MLT1G1"

  Query Match             95.0%;  Score 19;  DB 9;  Length 179212;
  Best Local Similarity   100.0%;  Pred. No. 96;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 ATTAGCATAATAAAATCTC 20
           ||||||||||||||||||
Db   10996 ATTAGCATAATAAAATCTC 10978


RESULT 14
AE008700/c
LOCUS       AE008700               22286 bp    DNA     linear   BCT 25-OCT-2001
DEFINITION  Salmonella typhimurium LT2, section 8 of 224 of the complete
            genome.
ACCESSION   AE008700 AE006468
VERSION     AE008700.1  GI:16418628
KEYWORDS    .
SOURCE      Salmonella typhimurium LT2.
  ORGANISM  Salmonella typhimurium LT2
            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
            Salmonella.
REFERENCE   1  (bases 1 to 22286)
  AUTHORS   McClelland,M., Sanderson,K.E., Spieth,J., Clifton,S.W.,
            Latreille,P., Courtney,L., Porwollik,S., Ali,J., Dante,M., Du,F.,
            Hou,S., Layman,D., Leonard,S., Nguyen,C., Scott,K., Holmes,A.,
            Grewal,N., Mulvaney,E., Ryan,E., Sun,H., Florea,L., Miller,W.,
            Stoneking,T., Nhan,M., Waterston,R. and Wilson,R.K.
  TITLE     Complete genome sequence of Salmonella enterica serovar Typhimurium
            LT2
  JOURNAL   Nature 413 (6858), 852-856 (2001)
   PUBMED   11677609
REFERENCE   2  (bases 1 to 22286)
  AUTHORS   The Salmonella typhimurium Genome Sequencing Project.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-MAR-2001) Genome Sequencing Center, Department of
            Genetics, Washington University School of Medicine, 4444 Forest
            Park Boulevard, St. Louis, MO 63108, USA
COMMENT     Supported by NIH grant 5U 01 AI43283

            Coding sequences below are predicted from manually evaluated
            computer analysis, using similarity information and the programs;
            GLIMMER; http://www.tigr.org/softlab/glimmer/glimmer.html and
            GeneMark; http://opal.biology.gatech.edu/GeneMark/

            EC numbers were kindly provided by Junko Yabuzaki and the Kyoto
            Encyclopedia of Genes and Genomes; http://www.genome.ad.jp/kegg/,
            and Pedro Romero and Peter Karp at EcoCyc;
            http://ecocyc.PangeaSystems.com/ecocyc/

            The analyses of ribosome binding sites and promoter binding sites
            were kindly provided by Heladia Salgado, Julio Collado-Vides and
            ReguonDB;
            http://kinich.cifn.unam.mx:8850/db/regulondb_intro.frameset

            This sequence was finished as follows unless otherwise noted: all
            regions were double stranded, sequenced with an alternate
            chemistries or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one m13 subclone.
FEATURES             Location/Qualifiers
     source          1. .22286
                     /organism="Salmonella typhimurium LT2"
                     /strain="LT2; SGSC 1412; ATCC 700720"
                     /db_xref="ATCC:700720"
                     /db_xref="taxon:99287"
     gene            84. .1571
                     /gene="murC"
                     /note="STM0129"
     RBS             84. .89
                     /gene="murC"
                     /note="Putative RBS for murC; RegulonDB:STMS1H000509"
     CDS             96. .1571
                     /gene="murC"
                     /EC_number="6.3.2.8"
                     /note="Ortholog of E. coli L-alanine adding enzyme,
                     UDP-N-acetyl-muramate:alanine ligase (AAC73202.1); Blastp
                     hit to AAC73202.1 (491 aa), 94% identity in aa 1 - 490"
                     /codon_start=1
                     /transl_table=11
                     /product="L-alanine adding enzyme,
                     UDP-N-acetyl-muramate:alanine ligase"
                     /protein_id="AAL19093.1"
                     /db_xref="GI:16418629"
                     /translation="MNTQQLAKLRSIVPEMRRVRHIHFVGIGGAGMGGIAEVLANEGY
                     QISGSDLAPNPVTQQLTSLGATIFFNHRPENVRDASVVVVSSAISADNPEIVAAHEAR
                     IPVIRRAEMLAELMRFRHGIAIAGTHGKTTTTAMVSSIYAEAGLDPTFVNGGLVKAAG
                     VHARLGHSRYLIAEADESDASFLHLQPMVAIVTNIEADHMDTYHGDFENLKQTFINFL
                     HNLPFYGRAVMCVDDPVIRELLPRVGRQTTTYGFSEDADVRVEDYQQIGPQGHFTLLR
                     QGMPDLHVTLNAPGRHNALNAAAAVAVATEEGIADDAILRALESFQGTGRRFDFLGEF
                     PLEPVNGKAGTAMLVDDYGHHPTEVDATIKAARAGWPDKNLVMLFQPHRYTRTRDLYD
                     DFANVLTQVDALLMLDVYPAGEAPIPGADSRSLCRTIRNRGKIDPILVSDPAQVATML
                     APVLTGNDLILVQGAGNVGKIARYLSEIKLKPQIQEEEQHG"
     RBS             1548. .1553
                     /gene="ddlB"
                     /note="Putative RBS for ddlB; RegulonDB:STMS1H000510"
     gene            1548. .2484
                     /gene="ddlB"
```

```
                /note="STM0130"
     CDS        1564. .2484
                /gene="ddlB"
                /EC_number="6.3.2.4"
                /note="Ortholog of E. coli D-alanine-D-alanine ligase B,
                affects cell division (AAC73203.1); Blastp hit to
                AAC73203.1 (306 aa), 90% identity in aa 1 - 306"
                /codon_start=1
                /transl_table=11
                /product="D-alanine-D-alanine ligase B, affects cell
                division"
                /protein_id="AAL19094.1"
                /db_xref="GI:16418630"
                /translation="MADKIAVLLGGTSAERDVSLNSGAAVLAGLREGGINAHPVDPQE
                VDVAQLKAMGFQKVFIALHGRGGEDGTLQGMLELLGLPYTGSGVMASALSMDKLRSKL
                LWQGAGLPVAPWVALTRAEFEKGLSEEQKARISALGLPLIVKPSREGSSVGMTKVVEE
                NALQGALSLAFQHDDEILIEKWLCGPEFTVAIVGEEILPSIRIQPAGTFYDYEAKYLS
                DETQYFCPAGLDASQEAALQSLVLQAWKALGCTGWGRIDVMLDSDGQFYLLEANTSPG
                MTSHSLVPMAARQAGMSFSQLVVRILELAD"
     gene       2328. .3316
                /gene="ftsQ"
                /note="STM0131"
     -35_signal 2328. .2336
                /gene="ftsQ"
                /note="Putative -35_signal for ftsQ;
                RegulonDB:STMLTH004481"
     -10_signal 2354. .2362
                /gene="ftsQ"
                /note="Putative -10_signal for ftsQ;
                RegulonDB:STMLTH004481"
     RBS        2467. .2472
                /gene="ftsQ"
                /note="Putative RBS for ftsQ; RegulonDB:STMS1H000511"
     CDS        2486. .3316
                /gene="ftsQ"
                /note="Ortholog of E. coli cell division protein; ingrowth
                of wall at septum (AAC73204.1); Blastp hit to AAC73204.1
                (276 aa), 93% identity in aa 1 - 276"
                /codon_start=1
                /transl_table=11
                /product="cell division protein; ingrowth of wall at
                septum"
                /protein_id="AAL19095.1"
                /db_xref="GI:16418631"
                /translation="MSQAALNTRNSDEETPSSRRNNGTRLAGILFLLTVLCTVFVSGW
                VVLGWMEDAQRLPLSKLVLTGERHYTRNDDIRQAILALGAPGTFMTQDVNIIQSQIER
                LPWIKQASVRKQWPDELKIHLVEYVPIARWNDQHMVDAEGNTFSVPSDRIGKQVLPML
                YGPEGSASEVLQGYREMGQVLAKDKFTLKEAAMTARRSWQLTLNNGIKLNLGRGDTMK
                RLARFVELYPVLQQQAQTDGKRISYVDLRYDSGAAVGWAPLPPEESNQQQNQAQAEQQ
                "
     gene       3299. .4575
                /gene="ftsA"
                /note="STM0132"
     RBS        3299. .3304
                /gene="ftsA"
                /note="Putative RBS for ftsA; RegulonDB:STMS1H000513"
     CDS        3313. .4575
                /gene="ftsA"
                /note="Ortholog of E. coli ATP-binding cell division
                protein, septation process, complexes with FtsZ,
                associated with junctions of inner and outer membranes
                (AAC73205.1); Blastp hit to AAC73205.1 (420 aa), 99%
                identity in aa 1 - 420"
                /codon_start=1
                /transl_table=11
                /product="ATP-binding cell division protein, septation
                process, complexes with FtsZ, associated with junctions of
                inner and outer membranes"
                /protein_id="AAL19096.1"
                /db_xref="GI:16418632"
                /translation="MIKATDRKLVVGLEIGTAKVAALVGEVLPDGMVNIIGVGSCPSR
                GMDKGGVNDLESVVKCVQRAIDQAELMADCQISSVYLALSGKHISCQNEIGMVPISEE
                EVTQEDVENVVHTAKSVRVRDEHRVLHVIPQEYAIDYQEGIKNPVGLSGVRMQAKVHL
                ITCHNDMAKNIVKAVERCGLKVDQLIFAGLAASYSVLTEDERELGVCVVDIGGGTMDI
                AVYTGGALRHTKVIPYAGNVVTSDIAYAFGTPPSDAEAIKVRHGCALGSIVGKDESVE
                VPSVGGRPPRSLQRQTLAEVIEPRYTELLNLVNEEILQLQEQLRQQGVKHHLAAGIVL
                TGGAAQIEGLAACAQRVFHTQVRIGAPLNITGLTDYAQEPYYSTAVGLLHYGKESHLN
                GEAEVEKRVTASVGSWIKRLNSWLRKEF"
     RBS        4623. .4628
                /gene="ftsZ"
                /note="Putative RBS for ftsZ; RegulonDB:STMS1H000514"
     gene       4623. .5787
                /gene="ftsZ"
                /note="STM0133"
     CDS        4636. .5787
                /gene="ftsZ"
                /note="Ortholog of E. coli cell division; forms
                circumferential ring; tubulin-like GTP-binding protein and
                GTPase (AAC73206.1); Blastp hit to AAC73206.1 (383 aa),
                98% identity in aa 1 - 383"
                /codon_start=1
                /transl_table=11
                /product="tubulin-like GTP-binding protein and GTPase,
                forms circumferential ring in cell division"
                /protein_id="AAL19097.1"
                /db_xref="GI:16418633"
                /translation="MFEPMELTNDAVIKVIGVGGGGGNAVEHMVRERIEGVEFFAVNT
                DAQALRKTAVGQTIQIGSGITKGLGAGANPEVGRNAADEDREALRAALEGADMVFIAA
                GMGGGTGTGAAPVVAEVAKDLGILTVAVVTKPFNFEGKKRMAFAEQGITELSKHVDSL
                ITIPNDKLLKVLGRGISLLDAFGAANDVLKGAVQGIAELITRPGLMNVDFADVRTVMS
                EMGYAMMGSGVASGEDRAEEAAEMAISSPLLEDIDLSGARGVLVNITAGFDLRLDEFE
                TVGNTIRAFASDNATVVIGTSLDPDMNDELRVTVVATGIGMDKRPEITLVTNKQVQQP
                VLDRYQQHGMAPLTQEQKTVAKVVNDNTPQAAKEPDYLDIPAFLRKQAD"
     RBS        5872. .5877
                /gene="lpxC"
                /note="Putative RBS for lpxC; RegulonDB:STMS1H000515"
     gene       5872. .6805
                /gene="lpxC"
                /note="STM0134"
     CDS        5888. .6805
                /gene="lpxC"
                /EC_number="3.5.1.-"
                /note="Ortholog of E. coli UDP-3-O-acyl
```

Query Match 92.0%; Score 18.4; DB 1; Length 22286;
Best Local Similarity 95.0%; Pred. No. 2.6e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 GATTAGCATAATAAAATCTC 20
          |||||||||||||| ||||
Db   7650 GATTAGCATAATAAAGTCTC 7631
```

```
RESULT 15
AC036102/c
LOCUS       AC036102              127603 bp    DNA     linear   PRI 16-OCT-2001
DEFINITION  Homo sapiens chromosome 15 clone CTD-2297L20 map 15q21.1, complete
            sequence.
ACCESSION   AC036102
VERSION     AC036102.8  GI:16152267
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 127603)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
            Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
            Pate,D. and Hood,L.
  TITLE     Sequencing of human chromosome 15 D15S146-D15S117 region
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 127603)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Abbasi,N., Baradarani,L., Birditt,B.,
            Bloom,S., Dors,M., Dickhoff,R., Fleetwood,P., Harrison,G.,
            James,R., Kaur,A., Madan,A., Owen,M.P., Ratcliffe,A., Shaffer,T.
            and Hood,L.
```

```
  TITLE     Direct Submission
  JOURNAL   Submitted (07-APR-2000) Multimegabase Sequencing Center, University
            of Washington, PO BOX 357730, Seattle, WA 98195, USA
REFERENCE   3  (bases 1 to 127603)
  AUTHORS   Rowen,L., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S.,
            Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R.,
            Pate,D. and Hood,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-OCT-2001) Multimegabase Sequencing Center, Institute
            for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA
            98105, USA
COMMENT     On Oct 16, 2001 this sequence version replaced gi:14318379.
            ------------ Genome Center
                Center: Multimegabase Sequencing Center
                Center code: UWMSC
                Web site: http://chroma.mbt.washington.edu/msg_www
                Contact: leerowen@systemsbiology.org
            ------------ Summary Statistics
                Sequencing vector: pUC18; L08752
                Chemistry: Dye-terminator Big Dye; 90% of reads
                Chemistry: Dye-primer Big Dye; 10% of reads
                Assembly program: Phrap; version 0.990399
                Note:data from AC013452 [Drafting center UWMSC] and AC022306
            [Drafting center UWMSC] were added for finishing.
FEATURES             Location/Qualifiers
     source          1. .127603
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="15"
                     /map="15q21.1"
                     /clone="CTD-2297L20"
                     /clone_lib="Cal Tech human BAC library 11"
                     /note="This clone overlaps RP11-325E5 AC013452 and
                     RP11-295H24 AC022306 . Data from overlapping BACs were
                     combined and the conensus sequence determined from
                     CTD-2297L20 to the extent possible."
     misc_feature    1. .113777
                     /note="overlap with RP11-325E5 AC013452 ."
     unsure          85642
                     /note="low quality data ."
     misc_feature    118220. .127603
                     /note="overlap with RP11-295H24 AC022306 ."
BASE COUNT    36642 a  23260 c  24122 g  43579 t
ORIGIN

  Query Match             92.0%;  Score 18.4;  DB 9;  Length 127603;
  Best Local Similarity   95.0%;  Pred. No. 1.8e+02;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GATTAGCATAATAAAATCTC 20
           || ||||||||||||||||||
Db   61460 GAGTAGCATAATAAAATCTC 61441


Search completed: October  6, 2002, 16:01:31
Job time : 2415 secs
```